

; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: All D peptides  
; OTHER INFORMATION: or peptidomimetics  
US-09-867-847-2

Query Match 100.0%; Score 40; DB 9; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | | | | |  
Db 16 KLVFFAED 23

RESULT 47

US-09-988-842-3

; Sequence 3, Application US/09988842  
; Patent No. US20020143105A1  
; GENERAL INFORMATION:  
; APPLICANT: Johansson, Jan  
; TITLE OF INVENTION: DISCORDANT HELIX STABILIZATION FOR PREVENTION  
; TITLE OF INVENTION: OF AMYLOID FORMATION  
; FILE REFERENCE: 12125-002001  
; CURRENT APPLICATION NUMBER: US/09/988,842  
; CURRENT FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: US 60/251,662  
; PRIOR FILING DATE: 2000-12-06  
; PRIOR APPLICATION NUMBER: US 60/253,695  
; PRIOR FILING DATE: 2000-11-20  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetically generated peptide  
US-09-988-842-3

Query Match 100.0%; Score 40; DB 9; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | | | | |  
Db 16 KLVFFAED 23

RESULT 48

US-09-851-071-3

; Sequence 3, Application US/09851071  
; Patent No. US20020177550A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmidt, Anne Marie

; APPLICANT: Stern, David  
; TITLE OF INVENTION: A METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A  
SUBJECT  
; FILE REFERENCE: 0575/55424-Z/JPW/SHS/MVM  
; CURRENT APPLICATION NUMBER: US/09/851,071  
; CURRENT FILING DATE: 2001-05-08  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Human  
US-09-851-071-3

Query Match 100.0%; Score 40; DB 9; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 49

US-09-962-955C-36

; Sequence 36, Application US/09962955C  
; Publication No. US20030013648A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerardo M. Castillo  
; APPLICANT: Alan D. Snow  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrick M. Dwyer  
; STREET: ProteoTech, Inc, 1818 Westlake Avenue N, Suite 114  
; CITY: Seattle  
; STATE: WA (Washington)  
; COUNTRY: United States of America  
; ZIP: 98109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette - 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: WordPerfect 9  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/962,955C  
; FILING DATE: 24-September-2001  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/938,275  
; FILING DATE: 22-August-2001  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dwyer, Patrick M.  
; REGISTRATION NUMBER: 32,411  
; REFERENCE/DOCKET NUMBER: PROTEO.P03CI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 343-7074

; TELEFAX: (206) 343-7085  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; ORIGINAL SOURCE:  
; ORGANISM: MOUSE  
; FEATURE:  
; OTHER INFORMATION: Also referred to in the specification as "AB 1-40"  
US-09-962-955C-36

Query Match 100.0%; Score 40; DB 10; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 50

US-09-792-079-12

; Sequence 12, Application US/09792079  
; Publication No. US20030083277A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Kentucky Research Foundation  
; APPLICANT: Hersh, Louis B.  
; APPLICANT: Mukherjee, Atish  
; TITLE OF INVENTION: Use Of Insulin Degrading Enzyme (IDE) For The Treatment  
Of Alzheimer's  
; TITLE OF INVENTION: Disease Patients  
; FILE REFERENCE: 050229-0261  
; CURRENT APPLICATION NUMBER: US/09/792,079  
; CURRENT FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/184,826  
; PRIOR FILING DATE: 2000-02-24  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-792-079-12

Query Match 100.0%; Score 40; DB 10; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 51

US-10-007-779A-1

```

; Sequence 1, Application US/10007779A
; Publication No. US20020168753A1
;   GENERAL INFORMATION:
;       APPLICANT: Castillo, Gerardo and Snow, Alan
;       TITLE OF INVENTION: In Vitro Formation of Congophilic
;                               Maltese-Cross Amyloid Plaques to Identify Anti-
Plaques
;                               Therapeutics for the Treatment of Alzheimer's and
Prion Diseases
;       NUMBER OF SEQUENCES: 1
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: Patrick M. Dwyer
;           STREET: ProteoTech, Inc., 1818 Westlake Ave N, Suite 114
;           CITY: Seattle
;           STATE: WA (Washington)
;           COUNTRY: USA
;           ZIP: 98109
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: 3.5 inch diskette
;           COMPUTER: PC
;           OPERATING SYSTEM: Windows 98
;           SOFTWARE: WordPerfect 9
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/10/007,779A
;           FILING DATE: 28-Apr-2002
;           CLASSIFICATION: Unknown
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 09/267,795
;           FILING DATE: 12-March-1999
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Dwyer, Patrick M.
;           REGISTRATION NUMBER: 32,411
;           REFERENCE/DOCKET NUMBER: PROTEO.P08
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: (206) 343-7074
;           TELEFAX: (206) 343-7085
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 40 AMINO ACIDS
;           TYPE: AMINO ACID
;           STRANDEDNESS: <Unknown>
;           TOPOLOGY: LINEAR
;       MOLECULE TYPE: PROTEIN
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-007-779A-1

```

```

Query Match          100.0%;  Score 40;  DB 13;  Length 40;
Best Local Similarity 100.0%;  Pred. No. 0.23;
Matches      8;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 KLVFFAED 8
        |||||
Db      16 KLVFFAED 23

```

```

RESULT 52
US-10-051-496-4

```



; Sequence 4, Application US/10051496  
; Publication No. US20020182660A1  
; GENERAL INFORMATION:  
; APPLICANT: Kei-Lai L. Fong  
; TITLE OF INVENTION: N- and C-Terminus Specific Immunoassays for  
; Full Length Beta-Amyloid Peptide - Abeta(1-40),  
Abeta(1-39),  
; Abeta(1-41), Abeta(1-42) and Abeta (1-43)  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kei-Lai L. Fong  
; STREET: 1004 West 8th Avenue  
; CITY: King of Prussia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.50 inch, 1.44MB storage  
; COMPUTER: IBM PC Compatibles  
; OPERATING SYSTEM: Windows  
; SOFTWARE: MS No. US20020182660A1epad  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/051,496  
; FILING DATE: 18-Jan-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/784,854A  
; FILING DATE: 16-Feb-2001  
; APPLICATION NUMBER: 60/183,407  
; FILING DATE: 18-February-2000  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Koenig, C. Frederick III  
; REGISTRATION NUMBER: 29,662  
; REFERENCE/DOCKET NUMBER: PBI-PT001.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-6400  
; TELEFAX: (215) 568-6499  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 Amino Acid  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Protein  
; FEATURE:  
; NAME/KEY: Signal Sequence  
; LOCATION: 1-40  
; IDENTIFICATION METHOD: Similarity to other sequences, hydro-  
phobic  
; OTHER INFORMATION:  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:

; RELEVANT RESIDUES IN SEQ ID NO: 4:FROM 1-40  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-051-496-4

Query Match 100.0%; Score 40; DB 13; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 53

US-10-217-584-3

; Sequence 3, Application US/10217584  
; Publication No. US20030077261A1  
; GENERAL INFORMATION:  
; APPLICANT: Paris, Daniel  
; APPLICANT: Mullan, Michael  
; TITLE OF INVENTION: Modulation of Angiogenesis by A-Beta Peptides  
; FILE REFERENCE: USF-T161XC1  
; CURRENT APPLICATION NUMBER: US/10/217,584  
; CURRENT FILING DATE: 2002-08-12  
; PRIOR APPLICATION NUMBER: 60/311,656  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(40)  
; OTHER INFORMATION: A-beta 1-40 peptide  
US-10-217-584-3

Query Match 100.0%; Score 40; DB 14; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 54

US-10-169-580-1

; Sequence 1, Application US/10169580  
; Publication No. US20030100477A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS FOR SUPPRESSING B-AMYLOID  
PRODUCTION  
; FILE REFERENCE: Q70898  
; CURRENT APPLICATION NUMBER: US/10/169,580

; CURRENT FILING DATE: 2002-07-08  
; PRIOR APPLICATION NUMBER: 2000-131037  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: PCT/JP01/03555  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-169-580-1

Query Match 100.0%; Score 40; DB 14; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 55

US-10-143-534-3

; Sequence 3, Application US/10143534  
; Publication No. US20030105152A1  
; GENERAL INFORMATION:  
; APPLICANT: Ingram, Vernon M.  
; APPLICANT: Blanchard, Barbara J.  
; APPLICANT: Stockwell, Brent R.  
; TITLE OF INVENTION: TREATMENTS FOR NEUROTOXICITY IN ALZHEIMER'S DISEASE  
; FILE REFERENCE: M00656/70078  
; CURRENT APPLICATION NUMBER: US/10/143,534  
; CURRENT FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: US 10/051,663  
; PRIOR FILING DATE: 2002-01-18  
; PRIOR APPLICATION NUMBER: US 09/706,574  
; PRIOR FILING DATE: 2000-11-03  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 3  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Peptide  
US-10-143-534-3

Query Match 100.0%; Score 40; DB 14; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 56

US-10-190-548A-4

; Sequence 4, Application US/10190548A  
 ; Publication No. US20030109435A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Griswold Prenner, Irene  
 ; APPLICANT: Wright, Sarah  
 ; APPLICANT: Yednock, Theodore  
 ; APPLICANT: Rydel, Russell  
 ; TITLE OF INVENTION: Methods of Inhibiting Amyloid Toxicity  
 ; FILE REFERENCE: 08576.0030-00  
 ; CURRENT APPLICATION NUMBER: US/10/190,548A  
 ; CURRENT FILING DATE: 2002-12-09  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 40  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-10-190-548A-4

Query Match 100.0%; Score 40; DB 14; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 0.23;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 16 KLVFFAED 23

RESULT 57

US-10-051-663-3

; Sequence 3, Application US/10051663  
 ; Publication No. US20030114510A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ingram, Vernon M.  
 ; APPLICANT: Blanchard, Barbara J.  
 ; APPLICANT: Stockwell, Brent R.  
 ; TITLE OF INVENTION: TREATMENTS FOR NEUROTOXICITY IN ALZHEIMER'S DISEASE  
 ; FILE REFERENCE: M0656/7071  
 ; CURRENT APPLICATION NUMBER: US/10/051,663  
 ; CURRENT FILING DATE: 2002-01-18  
 ; PRIOR APPLICATION NUMBER: US 09/706,574  
 ; PRIOR FILING DATE: 2000-11-03  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn Version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 40  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic Peptide  
 US-10-051-663-3

Query Match 100.0%; Score 40; DB 14; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 0.23;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 58

US-10-151-614-1

; Sequence 1, Application US/10151614  
; Publication No. US20030147811A1  
; GENERAL INFORMATION:  
; APPLICANT: WISNIEWSKI, Thomas  
; APPLICANT: TURNBULL, Daniel  
; APPLICANT: SIGURDSSON, Einar  
; APPLICANT: ZAIM WADGHIRI, Youssef  
; TITLE OF INVENTION: DETECTION OF ALZHEIMER'S AMYLOID BY MAGNETIC RESONANCE  
; FILE REFERENCE: WISNIEWSKI 2A  
; CURRENT APPLICATION NUMBER: US/10/151,614  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/292,625  
; PRIOR FILING DATE: 2001-05-23  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-151-614-1

Query Match 100.0%; Score 40; DB 14; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 59

US-10-159-279-12

; Sequence 12, Application US/10159279  
; Publication No. US20030165481A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Kentucky Research Foundation  
; APPLICANT: Hersh, Louis B.  
; APPLICANT: Mukherjee, Atish  
; TITLE OF INVENTION: Use Of Insulin Degrading Enzyme (IDE) For The Treatment Of Alzheimer's  
; TITLE OF INVENTION: Disease Patients  
; FILE REFERENCE: 050229-0298  
; CURRENT APPLICATION NUMBER: US/10/159,279  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 60/184,826  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 09/792,079  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-159-279-12

Query Match 100.0%; Score 40; DB 14; Length 40;  
Best Local Similarity 100.0%; Pred. No: 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 60

US-10-301-488A-7

; Sequence 7, Application US/10301488A  
; Publication No. US20030166558A1  
; GENERAL INFORMATION:  
; APPLICANT: FRANGIONE, Blas  
; APPLICANT: WISNIEWSKI, Thomas  
; APPLICANT: SIGURDSSON, Einar  
; TITLE OF INVENTION: SYNTHETIC IMMUNOGENIC BUT NON-DEPOSIT-FORMING  
POLYPEPTIDES AND  
; TITLE OF INVENTION: PEPTIDES HOMOLOGOUS TO AMYLOID BETA, PRION PROTEIN,  
AMYLIN,  
; TITLE OF INVENTION: ALPHA-SYNUCLEIN, OR POLYGLUTAMINE REPEATS FOR INDUCTION  
OF AN  
; TITLE OF INVENTION: IMMUNE RESPONSE THERETO  
; FILE REFERENCE: 5986/1K434US1  
; CURRENT APPLICATION NUMBER: US/10/301,488A  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/331,801  
; PRIOR FILING DATE: 2001-11-21  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid residues 1-6 can either be absent or present  
as Lys or  
; OTHER INFORMATION: Asp to form, in combination with residues 7-10, a N-  
terminal  
; OTHER INFORMATION: polylysine or polyaspartate segment of 4 to10 residues in  
length.  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: The C-terminal Ala residue may be amidated.  
US-10-301-488A-7

Query Match 100.0%; Score 40; DB 14; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 26 KLVFFAED 33

RESULT 61

US-10-301-488A-8

; Sequence 8, Application US/10301488A

; Publication No. US20030166558A1

; GENERAL INFORMATION:

; APPLICANT: FRANGIONE, Blas

; APPLICANT: WISNIEWSKI, Thomas

; APPLICANT: SIGURDSSON, Einar

; TITLE OF INVENTION: SYNTHETIC IMMUNOGENIC BUT NON-DEPOSIT-FORMING

POLYPEPTIDES AND

; TITLE OF INVENTION: PEPTIDES HOMOLOGOUS TO AMYLOID BETA, PRION PROTEIN,

AMYLIN,

; TITLE OF INVENTION: ALPHA-SYNUCLEIN, OR POLYGLUTAMINE REPEATS FOR INDUCTION  
OF AN

; TITLE OF INVENTION: IMMUNE RESPONSE THERETO

; FILE REFERENCE: 5986/1K434US1

; CURRENT APPLICATION NUMBER: US/10/301,488A

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 60/331,801

; PRIOR FILING DATE: 2001-11-21

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 40

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (31)..(40)

; OTHER INFORMATION: Amino acid residues 35-40 can either be absent or present  
as Lys

; OTHER INFORMATION: or Asp to form, in combination with residues 31-34, a C-  
terminal

; OTHER INFORMATION: polylysine or polyaspartate segment of 4-10 residues in  
length.

US-10-301-488A-8

Query Match 100.0%; Score 40; DB 14; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 62

US-10-366-125-27

```
; Sequence 27, Application US/10366125
; Publication No. US20030228259A1
; GENERAL INFORMATION:
; APPLICANT: Hellerstein, Marc
; TITLE OF INVENTION: MEASUREMENT OF BIOSYNTHESIS AND BREAKDOWN RATES OF
; TITLE OF INVENTION: BIOLOGICAL MOLECULES THAT ARE INACCESSIBLE OR NOT
; TITLE OF INVENTION: EASILY ACCESSIBLE TO DIRECT SAMPLING, NON-INVASIVELY,
; TITLE OF INVENTION: BY LABEL INCORPORATION INTO METABOLIC DERIVATIVES AND
; TITLE OF INVENTION: CATABOLITIC PRODUCTS
; FILE REFERENCE: 416272003500
; CURRENT APPLICATION NUMBER: US/10/366,125
; CURRENT FILING DATE: 2003-02-12
; PRIOR APPLICATION NUMBER: US 60/356,008
; PRIOR FILING DATE: 2002-02-12
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-366-125-27
```

```
Query Match          100.0%; Score 40; DB 15; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches      8; Conservative    0; Mismatches    0; Indels    0; Gaps    0;
```

```
Qy      1 KLVFFAED 8
        |||||
Db      16 KLVFFAED 23
```

RESULT 63

US-10-051-496-3

```
; Sequence 3, Application US/10051496
; Publication No. US20020182660A1
; GENERAL INFORMATION:
; APPLICANT: Kei-Lai L. Fong
; TITLE OF INVENTION: N- and C-Terminus Specific Immunoassays for
; Full Length Beta-Amyloid Peptide - Abeta(1-40),
Abeta(1-39),
; Abeta(1-41), Abeta(1-42) and Abeta (1-43)
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kei-Lai L. Fong
; STREET: 1004 West 8th Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 inch, 1.44MB storage
; COMPUTER: IBM PC Compatibles
; OPERATING SYSTEM: Windows
; SOFTWARE: MS No. US20020182660A1epad
```



```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/051,496
; FILING DATE: 18-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/784,854A
; FILING DATE: 16-Feb-2001
; APPLICATION NUMBER: 60/183,407
; FILING DATE: 18-February-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Koenig, C. Frederick III
; REGISTRATION NUMBER: 29,662
; REFERENCE/DOCKET NUMBER: PBI-PT001.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-6400
; TELEFAX: (215) 568-6499
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 Amino Acid
; TYPE: Amino Acid
; TOPOLOGY: Linear
; MOLECULE TYPE: Protein
; FEATURE:
; NAME/KEY: Signal Sequence
; LOCATION: 1-41
; IDENTIFICATION METHOD: Similarity to other sequences, hydro-
phobic
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1-41
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-051-496-3

```

```

Query Match          100.0%; Score 40; DB 13; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches      8; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

Qy      1 KLVFFAED 8
        |||||
Db      16 KLVFFAED 23

```

#### RESULT 64

US-10-190-548A-3

```

; Sequence 3, Application US/10190548A
; Publication No. US20030109435A1
; GENERAL INFORMATION:
; APPLICANT: Griswold Prenner, Irene
; APPLICANT: Wright, Sarah
; APPLICANT: Yednock, Theodore

```

; APPLICANT: Rydel, Russell  
; TITLE OF INVENTION: Methods of Inhibiting Amyloid Toxicity  
; FILE REFERENCE: 08576.0030-00  
; CURRENT APPLICATION NUMBER: US/10/190,548A  
; CURRENT FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-190-548A-3

Query Match 100.0%; Score 40; DB 14; Length 41;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | | | | |  
Db 16 KLVFFAED 23

RESULT 65

US-08-923-055-2

; Sequence 2, Application US/08923055  
; Publication No. US20010016327A1  
; GENERAL INFORMATION:  
; APPLICANT: Dana Giulian  
; TITLE OF INVENTION: Identification of Agents that Protect  
; TITLE OF INVENTION: Against Inflammatory Injury to Neurons  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz  
; ADDRESSEE: & No. US20010016327A1ris LLP  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: WORDPERFECT for WINDOWS 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/923,055  
; FILING DATE: Sept-03-97  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lori Y. Beardell  
; REGISTRATION NUMBER: 34,293  
; REFERENCE/DOCKET NUMBER: BYLR-0038  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100

; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-923-055-2

Query Match 100.0%; Score 40; DB 8; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 66

US-09-867-847-1

; Sequence 1, Application US/09867847  
; Patent No. US20020094335A1  
; GENERAL INFORMATION:  
; APPLICANT: Chalifour, Robert  
; APPLICANT: Hebert, Lise  
; APPLICANT: Kong, Xianqi  
; APPLICANT: Gervais, Francine  
; TITLE OF INVENTION: VACCINE FOR THE PREVENTION AND TREATMENT OF ALZHEIMER'S  
; TITLE OF INVENTION: AND AMYLOID RELATED DISEASES  
; FILE REFERENCE: 14445-501 CIP  
; CURRENT APPLICATION NUMBER: US/09/867,847  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: 60/168,594  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: 09/724,842  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: All D peptides  
; OTHER INFORMATION: or peptidomimetics  
US-09-867-847-1

Query Match 100.0%; Score 40; DB 9; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 67

US-09-956-625-26  
 ; Sequence 26, Application US/09956625  
 ; Patent No. US20020119926A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fraser, Paul  
 ; TITLE OF INVENTION: Inhibitors of IAPP Fibril Formation and Uses Thereof  
 ; FILE REFERENCE: 14445-503  
 ; CURRENT APPLICATION NUMBER: US/09/956,625  
 ; CURRENT FILING DATE: 2001-09-19  
 ; PRIOR APPLICATION NUMBER: 60/233,482  
 ; PRIOR FILING DATE: 2000-09-19  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 26  
 ; LENGTH: 42  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-956-625-26

Query Match 100.0%; Score 40; DB 9; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 16 KLVFFAED 23

RESULT 68  
 US-09-731-460-1  
 ; Sequence 1, Application US/09731460  
 ; Patent No. US20020137112A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chojkier, Mario  
 ; APPLICANT: Buck, Martina  
 ; TITLE OF INVENTION: Compositions and Methods for Diagnosing Alzheimer's  
 ; TITLE OF INVENTION: Disease  
 ; FILE REFERENCE: CHOJKIER-04302  
 ; CURRENT APPLICATION NUMBER: US/09/731,460  
 ; CURRENT FILING DATE: 2000-12-07  
 ; NUMBER OF SEQ ID NOS: 1  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 42  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 US-09-731-460-1

Query Match 100.0%; Score 40; DB 9; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 16 KLVFFAED 23

RESULT 69

US-09-962-955C-37

; Sequence 37, Application US/09962955C

; Publication No. US20030013648A1

; GENERAL INFORMATION:

; APPLICANT: Gerardo M. Castillo

; APPLICANT: Alan D. Snow

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Patrick M. Dwyer

; STREET: ProteoTech, Inc, 1818 Westlake Avenue N, Suite 114

; CITY: Seattle

; STATE: WA (Washington)

; COUNTRY: United States of America

; ZIP: 98109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette - 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM PC

; OPERATING SYSTEM: Windows 98

; SOFTWARE: WordPerfect 9

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/962,955C

; FILING DATE: 24-September-2001

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/938,275

; FILING DATE: 22-August-2001

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Dwyer, Patrick M.

; REGISTRATION NUMBER: 32,411

; REFERENCE/DOCKET NUMBER: PROTEO.P03CI

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 343-7074

; TELEFAX: (206) 343-7085

; INFORMATION FOR SEQ ID NO: 37:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 42 AMINO ACIDS

; TYPE: AMINO ACID

; STRANDEDNESS:

; TOPOLOGY: LINEAR

; ORIGINAL SOURCE:

; ORGANISM: MOUSE

; FEATURE:

; OTHER INFORMATION: Also referred to in the specification as "AB 1-42"

US-09-962-955C-37

Query Match 100.0%; Score 40; DB 10; Length 42;

Best Local Similarity 100.0%; Pred. No. 0.24;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8

|||||||

Db 16 KLVFFAED 23

RESULT 70

US-09-848-616-174

; Sequence 174, Application US/09848616  
; Publication No. US20030054010A1  
; GENERAL INFORMATION:  
; APPLICANT: Sebbel, Peter  
; APPLICANT: Dunant, Nicolas  
; APPLICANT: Bachmann, Martin  
; APPLICANT: Tissot, Alain  
; APPLICANT: Lechner, Franziska  
; TITLE OF INVENTION: Molecular Antigen Array  
; FILE REFERENCE: 1700.0180002  
; CURRENT APPLICATION NUMBER: US/09/848,616  
; CURRENT FILING DATE: 2001-05-05  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 174  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Amyloid Beta Peptide  
US-09-848-616-174

Query Match 100.0%; Score 40; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 71

US-09-865-294-65

; Sequence 65, Application US/09865294  
; Publication No. US20030068325A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: Immunogenic peptide composition as vaccines for the  
; TITLE OF INVENTION: prevention and treatment of Alzheimer's Disease  
; FILE REFERENCE: 1151-4167  
; CURRENT APPLICATION NUMBER: US/09/865,294  
; CURRENT FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 65  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-865-294-65

Query Match 100.0%; Score 40; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 72

US-09-792-079-13

; Sequence 13, Application US/09792079  
; Publication No. US20030083277A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Kentucky Research Foundation  
; APPLICANT: Hersh, Louis B.  
; APPLICANT: Mukherjee, Atish  
; TITLE OF INVENTION: Use Of Insulin Degrading Enzyme (IDE) For The Treatment  
Of Alzheimer's  
; TITLE OF INVENTION: Disease Patients  
; FILE REFERENCE: 050229-0261  
; CURRENT APPLICATION NUMBER: US/09/792,079  
; CURRENT FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/184,826  
; PRIOR FILING DATE: 2000-02-24  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-792-079-13

Query Match 100.0%; Score 40; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 73

US-09-825-242-1

; Sequence 1, Application US/09825242  
; Publication No. US20030092000A1  
; GENERAL INFORMATION:  
; APPLICANT: Schenk, Dale B.  
; APPLICANT: Neuralab Limited  
; TITLE OF INVENTION: Prevention and Treatment of Amyloidogenic Disease  
; FILE REFERENCE: 15270J-004720US  
; CURRENT APPLICATION NUMBER: US/09/825,242  
; CURRENT FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 09/201,430  
; PRIOR FILING DATE: 1998-11-30  
; PRIOR APPLICATION NUMBER: US 60/080,970  
; PRIOR FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 42

; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human Abeta42 beta-amyloid peptide  
US-09-825-242-1

Query Match 100.0%; Score 40; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 74

US-09-930-915A-293  
; Sequence 293, Application US/09930915A  
; Publication No. US20030138769A1  
; GENERAL INFORMATION:  
; APPLICANT: Birkett, Ashley J.  
; TITLE OF INVENTION: IMMUNOGENIC HBc CHIMER PARTICLES HAVING ENHANCED  
; TITLE OF INVENTION: STABILITY  
; FILE REFERENCE: 4564/83501 ICC-102.2 PCT  
; CURRENT APPLICATION NUMBER: US/09/930,915A  
; CURRENT FILING DATE: 2001-08-15  
; PRIOR APPLICATION NUMBER: 60/226,867  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,843  
; PRIOR FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 313  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 293  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-930-915A-293

Query Match 100.0%; Score 40; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 75

US-10-051-496-2  
; Sequence 2, Application US/10051496  
; Publication No. US20020182660A1  
; GENERAL INFORMATION:  
; APPLICANT: Kei-Lai L. Fong  
; TITLE OF INVENTION: N- and C-Terminus Specific Immunoassays for  
; Full Length Beta-Amyloid Peptide - Abeta(1-40),  
Abeta(1-39),  
; Abeta(1-41), Abeta(1-42) and Abeta (1-43)



```

;      NUMBER OF SEQUENCES: 5
;      CORRESPONDENCE ADDRESS:
;          ADDRESSEE: Kei-Lai L. Fong
;          STREET: 1004 West 8th Avenue
;          CITY: King of Prussia
;          STATE: Pennsylvania
;          COUNTRY: USA
;          ZIP: 19406
;      COMPUTER READABLE FORM:
;          MEDIUM TYPE: 3.50 inch, 1.44MB storage
;          COMPUTER: IBM PC Compatibles
;          OPERATING SYSTEM: Windows
;          SOFTWARE: MS No. US20020182660A1epad
;      CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/10/051,496
;          FILING DATE: 18-Jan-2002
;          CLASSIFICATION: <Unknown>
;      PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: US/09/784,854A
;          FILING DATE: 16-Feb-2001
;          APPLICATION NUMBER: 60/183,407
;          FILING DATE: 18-February-2000
;      ATTORNEY/AGENT INFORMATION:
;          NAME: Koenig, C. Frederick III
;          REGISTRATION NUMBER: 29,662
;          REFERENCE/DOCKET NUMBER: FBI-PT001.1
;      TELECOMMUNICATION INFORMATION:
;          TELEPHONE: (215) 568-6400
;          TELEFAX: (215) 568-6499
;      INFORMATION FOR SEQ ID NO: 2:
;          SEQUENCE CHARACTERISTICS:
;              LENGTH: 42 Amino Acid
;              TYPE: Amino Acid
;              TOPOLOGY: Linear
;          MOLECULE TYPE: Protein
;          FEATURE:
;              NAME/KEY:    Signal Sequence
;              LOCATION:    1-42
;              IDENTIFICATION METHOD:    Similarity to other sequences, hydro-
phobic
;          OTHER INFORMATION:
;      PUBLICATION INFORMATION:
;          AUTHORS:
;          TITLE:
;          JOURNAL:
;          VOLUME:
;          ISSUE:
;          PAGES:
;          DATE:
;          RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1-42
;      SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-051-496-2

```

```

Query Match          100.0%;  Score 40;  DB 13;  Length 42;
Best Local Similarity 100.0%;  Pred. No. 0.24;
Matches      8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

```

Qy           1 KLVFFAED 8  
              |||||||  
Db           16 KLVFFAED 23

RESULT 76

US-10-082-804-7

; Sequence 7, Application US/10082804  
; Publication No. US20020194632A1  
; GENERAL INFORMATION:  
; APPLICANT: McConlogue, Lisa  
; APPLICANT: Gurney, Mark E.  
; TITLE OF INVENTION: Transgenic Knockouts of BACE-1  
; FILE REFERENCE: MBHB 02-329-A  
; CURRENT APPLICATION NUMBER: US/10/082,804  
; CURRENT FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: 60/271,092  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 60/271,514  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/293,762  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: A-beta 42 sequence.  
US-10-082-804-7

Query Match                   100.0%; Score 40; DB 13; Length 42;  
Best Local Similarity   100.0%; Pred. No. 0.24;  
Matches       8; Conservative   0; Mismatches   0; Indels    0; Gaps    0;

Qy           1 KLVFFAED 8  
              |||||||  
Db           16 KLVFFAED 23

RESULT 77

US-10-217-584-2

; Sequence 2, Application US/10217584  
; Publication No. US20030077261A1  
; GENERAL INFORMATION:  
; APPLICANT: Paris, Daniel  
; APPLICANT: Mullan, Michael  
; TITLE OF INVENTION: Modulation of Angiogenesis by A-Beta Peptides  
; FILE REFERENCE: USF-T161XC1  
; CURRENT APPLICATION NUMBER: US/10/217,584  
; CURRENT FILING DATE: 2002-08-12  
; PRIOR APPLICATION NUMBER: 60/311,656  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2

; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(42)  
; OTHER INFORMATION: A-beta 1-42 peptide  
US-10-217-584-2

Query Match 100.0%; Score 40; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 78

US-10-169-580-2

; Sequence 2, Application US/10169580  
; Publication No. US20030100477A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS FOR SUPPRESSING B-AMYLOID  
PRODUCTION  
; FILE REFERENCE: Q70898  
; CURRENT APPLICATION NUMBER: US/10/169,580  
; CURRENT FILING DATE: 2002-07-08  
; PRIOR APPLICATION NUMBER: 2000-131037  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: PCT/JP01/03555  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-169-580-2

Query Match 100.0%; Score 40; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 79

US-10-278-181-1

; Sequence 1, Application US/10278181  
; Publication No. US20030104488A1  
; GENERAL INFORMATION:  
; APPLICANT: Chojkier, Mario  
; APPLICANT: Buck, Martina

```
; TITLE OF INVENTION: Compositions and Methods for Diagnosing Alzheimer's
; TITLE OF INVENTION: Disease
; FILE REFERENCE: CHOJKIER-04302
; CURRENT APPLICATION NUMBER: US/10/278,181
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/731,460
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-278-181-1
```

```
Query Match          100.0%; Score 40; DB 14; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches      8; Conservative    0; Mismatches    0; Indels      0; Gaps      0;
```

```
Qy          1 KLVFFAED 8
             |||||
Db          16 KLVFFAED 23
```

# RESULT 80

```
US-10-143-534-2
; Sequence 2, Application US/10143534
; Publication No. US20030105152A1
; GENERAL INFORMATION:
; APPLICANT: Ingram, Vernon M.
; APPLICANT: Blanchard, Barbara J.
; APPLICANT: Stockwell, Brent R.
; TITLE OF INVENTION: TREATMENTS FOR NEUROTOXICITY IN ALZHEIMER'S DISEASE
; FILE REFERENCE: M00656/70078
; CURRENT APPLICATION NUMBER: US/10/143,534
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: US 10/051,663
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 09/706,574
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Version 3.0
; SEQ ID NO 2
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-143-534-2
```

```
Query Match          100.0%; Score 40; DB 14; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches      8; Conservative    0; Mismatches    0; Indels      0; Gaps      0;
```

```
Qy          1 KLVFFAED 8
```

Db                   |||||||  
16 KLVFFAED 23

RESULT 81

US-10-190-548A-1  
; Sequence 1, Application US/10190548A  
; Publication No. US20030109435A1  
; GENERAL INFORMATION:  
; APPLICANT: Griswold Prenner, Irene  
; APPLICANT: Wright, Sarah  
; APPLICANT: Yednock, Theodore  
; APPLICANT: Rydel, Russell  
; TITLE OF INVENTION: Methods of Inhibiting Amyloid Toxicity  
; FILE REFERENCE: 08576.0030-00  
; CURRENT APPLICATION NUMBER: US/10/190,548A  
; CURRENT FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-190-548A-1

Query Match                   100.0%; Score 40; DB 14; Length 42;  
Best Local Similarity       100.0%; Pred. No. 0.24;  
Matches       8; Conservative       0; Mismatches       0; Indels       0; Gaps       0;

Qy                   1 KLVFFAED 8  
                     |||||||  
Db                   16 KLVFFAED 23

RESULT 82

US-10-051-663-2  
; Sequence 2, Application US/10051663  
; Publication No. US20030114510A1  
; GENERAL INFORMATION:  
; APPLICANT: Ingram, Vernon M.  
; APPLICANT: Blanchard, Barbara J.  
; APPLICANT: Stockwell, Brent R.  
; TITLE OF INVENTION: TREATMENTS FOR NEUROTOXICITY IN ALZHEIMER'S DISEASE  
; FILE REFERENCE: M0656/7071  
; CURRENT APPLICATION NUMBER: US/10/051,663  
; CURRENT FILING DATE: 2002-01-18  
; PRIOR APPLICATION NUMBER: US 09/706,574  
; PRIOR FILING DATE: 2000-11-03  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 2  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Peptide  
US-10-051-663-2

Query Match 100.0%; Score 40; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 83

US-10-159-279-13

; Sequence 13, Application US/10159279  
; Publication No. US20030165481A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Kentucky Research Foundation  
; APPLICANT: Hersh, Louis B.  
; APPLICANT: Mukherjee, Atish  
; TITLE OF INVENTION: Use Of Insulin Degrading Enzyme (IDE) For The Treatment  
Of Alzheimer's  
; TITLE OF INVENTION: Disease Patients  
; FILE REFERENCE: 050229-0298  
; CURRENT APPLICATION NUMBER: US/10/159,279  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: 60/184,826  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 09/792,079  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-159-279-13

Query Match 100.0%; Score 40; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 84

US-10-318-302-4

; Sequence 4, Application US/10318302  
; Publication No. US20030171556A1  
; GENERAL INFORMATION:  
; APPLICANT: POSCO  
; APPLICANT: POSTECH FOUNDATION  
; APPLICANT: Chae, Chi-Bom  
; APPLICANT: Gho, Yong Song  
; APPLICANT: Yang, Seung-Pil  
; APPLICANT: Kwon, Byung Oh  
; APPLICANT: Bae, Dong-Goo

; APPLICANT: Hwang, Sewook  
; TITLE OF INVENTION: BETA-AMYLOID BINDING FACTORS AND INHIBITORS THEREOF  
; FILE REFERENCE: 10011-00001  
; CURRENT APPLICATION NUMBER: US/10/318,302  
; CURRENT FILING DATE: 2002-12-12  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-318-302-4

Query Match 100.0%; Score 40; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 85

US-10-050-902-220  
; Sequence 220, Application US/10050902  
; Publication No. US20030175290A1  
; GENERAL INFORMATION:  
; APPLICANT: Renner, Wolfgang A.  
; APPLICANT: Bachmann, Martin  
; APPLICANT: Tissot, Alain  
; APPLICANT: Maurer, Patrick  
; APPLICANT: Lechner, Franziska  
; APPLICANT: Sebbel, Peter  
; APPLICANT: Piossek, Christine  
; TITLE OF INVENTION: Molecular Antigen Array  
; FILE REFERENCE: 1700.0190004  
; CURRENT APPLICATION NUMBER: US/10/050,902  
; CURRENT FILING DATE: 2002-01-18  
; PRIOR APPLICATION NUMBER: US 60/262,379  
; PRIOR FILING DATE: 2001-01-19  
; PRIOR APPLICATION NUMBER: US 60/288,549  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/326,998  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: US 60/331,045  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 220  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Amyloid Beta Peptide  
US-10-050-902-220

Query Match 100.0%; Score 40; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 86

US-10-050-898-220  
; Sequence 220, Application US/10050898  
; Publication No. US20030175711A1  
; GENERAL INFORMATION:  
; APPLICANT: Renner, Wolfgang A.  
; APPLICANT: Bachmann, Martin  
; APPLICANT: Tissot, Alain  
; APPLICANT: Maurer, Patrick  
; APPLICANT: Lechner, Franziska  
; APPLICANT: Sebbel, Peter  
; APPLICANT: Piossek, Christine  
; APPLICANT: Ortman, Rainer  
; APPLICANT: Luond, Rainer  
; APPLICANT: Staufenberg, Matthias  
; APPLICANT: Frey, Peter  
; TITLE OF INVENTION: Molecular Antigen Array  
; FILE REFERENCE: 1700.0190005  
; CURRENT APPLICATION NUMBER: US/10/050,898  
; CURRENT FILING DATE: 2002-01-18  
; PRIOR APPLICATION NUMBER: US 60/262,379  
; PRIOR FILING DATE: 2001-01-19  
; PRIOR APPLICATION NUMBER: US 60/288,549  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/326,998  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: US 60/331,045  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 220  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Amyloid Beta Peptide  
US-10-050-898-220

Query Match 100.0%; Score 40; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 87

US-10-082-014-81  
; Sequence 81, Application US/10082014  
; Publication No. US20030185858A1  
; GENERAL INFORMATION:  
; APPLICANT: Birkett, Ashley J.



```
; TITLE OF INVENTION: IMMUNOGENIC HBc CHIMER PARTICLES STABILIZED WITH AN N-
TERMINAL CYSTEINE
; FILE REFERENCE: ICC-130.0 4564/85124
; CURRENT APPLICATION NUMBER: US/10/082,014
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 09/930,915
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Alzheimer's disease b-Amyloid
US-10-082-014-81
```

```
Query Match          100.0%; Score 40; DB 14; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches      8; Conservative    0; Mismatches    0; Indels      0; Gaps      0;
```

```
Qy          1 KLVFFAED 8
             |||||
Db          16 KLVFFAED 23
```

```
RESULT 88
US-10-372-076-82
; Sequence 82, Application US/10372076
; Publication No. US20030198645A1
; GENERAL INFORMATION:
; APPLICANT: Page, Mark
; APPLICANT: Friede, Martin
; TITLE OF INVENTION: STABILIZED HBc CHIMER PARTICLES AS THERAPEUTIC VACCINE
FOR
; TITLE OF INVENTION: CHRONIC HEPATITIS
; FILE REFERENCE: 4564/87179
; CURRENT APPLICATION NUMBER: US/10/372,076
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: 10/080,299
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/082,014
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Alzheimer's disease b-Amyloid
US-10-372-076-82
```

```
Query Match          100.0%; Score 40; DB 14; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches      8; Conservative    0; Mismatches    0; Indels      0; Gaps      0;
```

```
Qy          1 KLVFFAED 8
             |||||
Db          16 KLVFFAED 23
```

RESULT 89

US-10-231-298B-15

```
; Sequence 15, Application US/10231298B
; Publication No. US20030219853A1
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Cross-Linking a Compound
; FILE REFERENCE: SAMG/0006
; CURRENT APPLICATION NUMBER: US/10/231,298B
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-231-298B-15
```

```
Query Match          100.0%;  Score 40;  DB 15;  Length 42;
Best Local Similarity 100.0%;  Pred. No. 0.24;
Matches      8;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;
```

```
Qy      1 KLVFFAED 8
        |||||
Db      16 KLVFFAED 23
```

RESULT 90

US-10-231-470C-15

```
; Sequence 15, Application US/10231470C
; Publication No. US20030219857A1
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method Of Producing Transglutaminase Having Broad
Substrate
; TITLE OF INVENTION: Activity
; FILE REFERENCE: SAMG/0003
; CURRENT APPLICATION NUMBER: US/10/231,470C
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-231-470C-15
```

```
Query Match          100.0%;  Score 40;  DB 15;  Length 42;
```

Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 91

US-10-366-125-28

; Sequence 28, Application US/10366125  
; Publication No. US20030228259A1  
; GENERAL INFORMATION:  
; APPLICANT: Hellerstein, Marc  
; TITLE OF INVENTION: MEASUREMENT OF BIOSYNTHESIS AND BREAKDOWN RATES OF  
; TITLE OF INVENTION: BIOLOGICAL MOLECULES THAT ARE INACCESSIBLE OR NOT  
; TITLE OF INVENTION: EASILY ACCESSIBLE TO DIRECT SAMPLING, NON-INVASIVELY,  
; TITLE OF INVENTION: BY LABEL INCORPORATION INTO METABOLIC DERIVATIVES AND  
; TITLE OF INVENTION: CATABOLITIC PRODUCTS  
; FILE REFERENCE: 416272003500  
; CURRENT APPLICATION NUMBER: US/10/366,125  
; CURRENT FILING DATE: 2003-02-12  
; PRIOR APPLICATION NUMBER: US 60/356,008  
; PRIOR FILING DATE: 2002-02-12  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-366-125-28

Query Match 100.0%; Score 40; DB 15; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 92

US-10-411-544-2

; Sequence 2, Application US/10411544  
; Publication No. US20030232758A1  
; GENERAL INFORMATION:  
; APPLICANT: St. George-Hyslop, Peter  
; APPLICANT: McLaurin, JoAnne  
; TITLE OF INVENTION: Immunological Methods and Compositions for the Treatment  
of Alzheimer's  
; TITLE OF INVENTION: Disease  
; FILE REFERENCE: LI01547  
; CURRENT APPLICATION NUMBER: US/10/411,544  
; CURRENT FILING DATE: 2003-04-10  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2

; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-411-544-2

Query Match 100.0%; Score 40; DB 15; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 93

US-10-231-213D-15  
; Sequence 15, Application US/10231213D  
; Publication No. US20040001848A1  
; GENERAL INFORMATION:  
; APPLICANT: Chou, Szu-Yi  
; TITLE OF INVENTION: Method of Producing Disease-Specific Antigens  
; FILE REFERENCE: SAMG/0005  
; CURRENT APPLICATION NUMBER: US/10/231,213D  
; CURRENT FILING DATE: 2002-08-28  
; PRIOR APPLICATION NUMBER: 60/361,166  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/363,445  
; PRIOR FILING DATE: 2002-03-08  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-231-213D-15

Query Match 100.0%; Score 40; DB 15; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 94

US-10-231-114C-15  
; Sequence 15, Application US/10231114C  
; Publication No. US20040005654A1  
; GENERAL INFORMATION:  
; APPLICANT: Chou, Szu-Yi  
; TITLE OF INVENTION: Method of Producing Polyvalent Antigens  
; FILE REFERENCE: SAMG/0002  
; CURRENT APPLICATION NUMBER: US/10/231,114C  
; CURRENT FILING DATE: 2002-08-28  
; PRIOR APPLICATION NUMBER: 60/361,166  
; PRIOR FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: 60/363,445  
; PRIOR FILING DATE: 2002-03-08  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-231-114C-15

Query Match 100.0%; Score 40; DB 15; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 95

US-09-280-966-1

; Sequence 1, Application US/09280966  
; Patent No. US20010020097A1  
; GENERAL INFORMATION:  
; APPLICANT: JAMES E. AUDIA  
; BEVERLY K. FOLMER  
; VARGHESE JOHN  
; JEFFREY S. NISSEN  
; WARREN J. PORTER  
; EUGENE D. THORSETT  
; JING WU  
; TITLE OF INVENTION: N-(ARYL/HETEROARYLACETYL) AMINO  
; ACID ESTERS, PHARMACEUTICAL COMPOSITIONS  
; COMPRISING SAME, AND METHODS FOR INHIBITING  
; -AMYLOID PEPTIDE RELEASE AND/OR ITS  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis, LLP  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/280,966  
; FILING DATE: 30-Mar-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/976,191  
; FILING DATE: 21 NOV 1997  
; APPLICATION NUMBER: 60/077,175  
; FILING DATE: 22 NOV 1996

```

;      ATTORNEY/AGENT INFORMATION:
;      NAME: Swiss, Gerald F.
;      REGISTRATION NUMBER: 30,113
;      REFERENCE/DOCKET NUMBER: 002010-335
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 650-622-2300
;      TELEFAX: 650-622-2499
;      INFORMATION FOR SEQ ID NO: 1:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 43 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: peptide
;      SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-280-966-1

```

```

Query Match          100.0%;  Score 40;  DB 9;  Length 43;
Best Local Similarity 100.0%;  Pred. No. 0.25;
Matches      8;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

Qy           1 KLVFFAED 8
             |||||
Db           16 KLVFFAED 23

```

```

RESULT 96
US-09-904-987-1
; Sequence 1, Application US/09904987
; Patent No. US20020037908A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020037908Alactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and
Prepathological
; TITLE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / QRHUA4
; DATABASE ENTRY DATE: 2000-09-15
; RELEVANT RESIDUES: (672)..(714)
US-09-904-987-1

```

```

Query Match          100.0%;  Score 40;  DB 9;  Length 43;
Best Local Similarity 100.0%;  Pred. No. 0.25;
Matches      8;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

Qy           1 KLVFFAED 8
             |||||
Db           16 KLVFFAED 23

```

RESULT 97

US-09-808-037-3

; Sequence 3, Application US/09808037  
; Patent No. US20020052311A1  
; GENERAL INFORMATION:  
; APPLICANT: SOLOMON, Beka  
; APPLICANT: HANAN, Eilat  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR  
DIAGNOSIS OF  
; TITLE OF INVENTION: NEUROLOGICAL DISEASES AND DISORDERS  
; FILE REFERENCE: SOLOMON=2D  
; CURRENT APPLICATION NUMBER: US/09/808,037  
; CURRENT FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 09/629,971  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 09/473,653  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: US 60/152,417  
; PRIOR FILING DATE: 1999-09-03  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-09-808-037-3

Query Match 100.0%; Score 40; DB 9; Length 43;  
Best Local Similarity 100.0%; Pred. No. 0.25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | | | | |  
Db 16 KLVFFAED 23

RESULT 98

US-09-866-712-3

; Sequence 3, Application US/09866712  
; Patent No. US20020058637A1  
; GENERAL INFORMATION:  
; APPLICANT: Akihiko TAKASHIMA et al.  
; TITLE OF INVENTION: METHODS OF INHIBITING TAU-PROTEIN KINASE I ACTIVITY  
; INHIBITING NEURONAL CELL DEATH AND TREATING  
ALZHEIMER'S DISEASE AND TAU-PROTEIN  
; KINASE I (AS AMENDED)  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WENDEROTH, LIND & PONACK  
; STREET: 2033 K Street, N.W., #800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20006

```

;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
;      COMPUTER: IBM Compatible
;      OPERATING SYSTEM: MS-DOS
;      SOFTWARE: Wordperfect 5.1
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/866,712
;      FILING DATE: 30-May-2001
;      CLASSIFICATION: <Unknown>
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 09/216,958
;      FILING DATE: December 21, 1998
;      ATTORNEY/AGENT INFORMATION:
;      NAME: Lee Cheng
;      REGISTRATION NUMBER: 40,949
;      REFERENCE/DOCKET NUMBER: 2001-0488/LC/00177
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (202)721-8200
;      TELEFAX: (202)721-8250
;      TELEX: <Unknown>
;      INFORMATION FOR SEQ ID NO: 3:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 43 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;      MOLECULE TYPE: peptide
;      SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

US-09-866-712-3

```

Query Match          100.0%;  Score 40;  DB 9;  Length 43;
Best Local Similarity 100.0%;  Pred. No. 0.25;
Matches      8;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 KLVFFAED 8
        |||||
Db      16 KLVFFAED 23

```

#### RESULT 99

US-09-972-475-1

```

; Sequence 1, Application US/09972475
; Patent No. US20020098173A1
;      GENERAL INFORMATION:
;      APPLICANT: Findeis, Mark A. et al.
;      TITLE OF INVENTION: Modulators of Amyloid Aggregation
;      NUMBER OF SEQUENCES: 45
;      CORRESPONDENCE ADDRESS:
;      ADDRESSEE: LAHIVE & COCKFIELD, LLP
;      STREET: 28 State Street
;      CITY: Boston
;      STATE: Massachusetts
;      COUNTRY: USA
;      ZIP: 02109-1875
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible

```



```

;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/972,475
;      FILING DATE: 04-Oct-2001
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 08/617,267
;      FILING DATE: <Unknown>
;      APPLICATION NUMBER: USSN 08/475,579
;      FILING DATE: 07-JUN-1995
;      APPLICATION NUMBER: USSN 08/548,998
;      FILING DATE: 27-OCT-1995
;      ATTORNEY/AGENT INFORMATION:
;      NAME: DeConti, Giulio A.
;      REGISTRATION NUMBER: 31,503
;      REFERENCE/DOCKET NUMBER: PPI-002CP2
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (617)227-7400
;      TELEFAX: (617)227-5941
;      INFORMATION FOR SEQ ID NO: 1:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 43 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: peptide
;      FRAGMENT TYPE: internal
;      SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-972-475-1

```

```

Query Match          100.0%;  Score 40;  DB 9;  Length 43;
Best Local Similarity 100.0%;  Pred. No. 0.25;
Matches      8;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 KLVFFAED 8
        |||||
Db      16 KLVFFAED 23

```

```

RESULT 100
US-09-992-800-1
; Sequence 1, Application US/09992800
; Patent No. US20020102261A1
; GENERAL INFORMATION:
; APPLICANT: Raso, Victor
; TITLE OF INVENTION: IMMUNOLOGICAL CONTROL OF BETA-AMYLOID LEVELS IN VIVO
; FILE REFERENCE: BBRI-2006
; CURRENT APPLICATION NUMBER: US/09/992,800
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 09/594,366
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/139,408
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT

```

; ORGANISM: Homo sapiens  
US-09-992-800-1

Query Match 100.0%; Score 40; DB 9; Length 43;  
Best Local Similarity 100.0%; Pred. No. 0.25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | | | | |  
Db 16 KLVFFAED 23

Search completed: February 28, 2004, 09:07:54  
Job time : 86.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 28, 2004, 08:37:49 ; Search time 27 Seconds  
(without alignments)  
28.501 Million cell updates/sec

Title: US-09-668-314C-73  
Perfect score: 40  
Sequence: 1 KLVFFAED 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	40	100.0	33	2	S23094	beta-amyloid prote
2	40	100.0	42	2	PN0512	beta-amyloid prote
3	40	100.0	57	2	E60045	Alzheimer's diseas
4	40	100.0	57	2	F60045	Alzheimer's diseas
5	40	100.0	57	2	G60045	Alzheimer's diseas
6	40	100.0	57	2	D60045	Alzheimer's diseas
7	40	100.0	57	2	A60045	Alzheimer's diseas
8	40	100.0	57	2	B60045	Alzheimer's diseas
9	40	100.0	82	2	PQ0438	Alzheimer's diseas
10	40	100.0	695	1	A49795	Alzheimer's diseas
11	40	100.0	695	2	A27485	Alzheimer's diseas
12	40	100.0	695	2	S00550	Alzheimer's diseas
13	40	100.0	770	1	QRHUA4	Alzheimer's diseas
14	36	90.0	747	2	JH0773	Alzheimer's diseas

15	34	85.0	321	2	H71729	hypothetical prote
16	32	80.0	182	2	T35807	hypothetical prote
17	32	80.0	261	2	B89868	conserved hypothet
18	31	77.5	119	2	D69345	LSU ribosomal prot
19	31	77.5	641	2	H69651	lichenan operon tr
20	31	77.5	1339	2	T38991	conserved hypothet
21	31	77.5	1364	2	T51920	probable xanthine
22	30	75.0	222	2	T24151	hypothetical prote
23	30	75.0	258	2	AG0459	Sec-independent pr
24	30	75.0	341	2	A64383	hypothetical prote
25	30	75.0	370	2	T47131	G-protein coupled
26	30	75.0	502	2	T27908	hypothetical prote
27	30	75.0	533	2	T46975	lysine-tRNA ligase
28	30	75.0	681	2	T39814	hypothetical prote
29	30	75.0	740	2	S61568	probable membrane
30	30	75.0	768	2	T45876	hypothetical prote
31	30	75.0	1353	2	JC4279	adenylate cyclase
32	29	72.5	99	2	F95064	ribosomal protein
33	29	72.5	99	2	H97931	conserved hypothet
34	29	72.5	100	2	AH1192	B. subtilis YneR p
35	29	72.5	116	1	R5HSS6	ribosomal protein
36	29	72.5	152	2	T06645	hypothetical prote
37	29	72.5	162	2	T13487	NADH2 dehydrogenas
38	29	72.5	162	2	T13563	NADH2 dehydrogenas
39	29	72.5	162	2	T13656	NADH2 dehydrogenas
40	29	72.5	162	2	T13659	NADH2 dehydrogenas
41	29	72.5	164	2	T13562	NADH2 dehydrogenas
42	29	72.5	189	2	S39864	late competence op
43	29	72.5	231	2	H85138	hypothetical prote
44	29	72.5	247	2	B86301	hypothetical prote
45	29	72.5	258	1	S39747	ywfN protein - Bac
46	29	72.5	273	2	T06661	hypothetical prote
47	29	72.5	286	2	H64054	site-specific DNA-
48	29	72.5	300	2	AC1342	ABC transporter (A
49	29	72.5	300	2	AI1712	ABC transporter (A
50	29	72.5	314	2	F86805	cation transporter
51	29	72.5	352	2	AH2097	sorbitol dehydroge
52	29	72.5	352	2	T48903	wax synthase [impo
53	29	72.5	354	2	E97128	magnesium and coba
54	29	72.5	357	2	S59678	HST2 protein - yea
55	29	72.5	369	2	A75091	phosphonoacetate h
56	29	72.5	428	2	T48008	hypothetical prote
57	29	72.5	469	2	A64422	conserved hypothet
58	29	72.5	508	2	T36945	hypothetical prote
59	29	72.5	590	2	T45820	hypothetical prote
60	29	72.5	607	2	T24172	hypothetical prote
61	29	72.5	621	2	A82153	PTS system, fructo
62	29	72.5	633	2	S61435	metalloproteinase
63	29	72.5	649	2	S42894	metalloproteinase
64	29	72.5	710	2	A96540	hypothetical prote
65	29	72.5	802	2	JH0595	potassium channel
66	29	72.5	853	1	CHRTD1	potassium channel
67	29	72.5	857	2	I56529	potassium channel
68	29	72.5	858	2	S31761	potassium channel
69	29	72.5	987	2	S62752	probable DNA-direc
70	29	72.5	1016	2	S30236	genome polyprotein
71	29	72.5	3488	2	T34418	hypothetical prote

72	29	72.5	4466	1	S17231	dynein beta heavy
73	29	72.5	4466	1	S17653	dynein beta heavy
74	28	70.0	58	2	B44358	cysteine-rich prot
75	28	70.0	96	2	A05035	translation initia
76	28	70.0	120	2	B64450	ribosomal protein
77	28	70.0	124	2	D71025	probable ribosomal
78	28	70.0	128	2	C75109	lsu ribosomal prot
79	28	70.0	130	2	S75397	ribosomal protein
80	28	70.0	176	2	S67150	hypothetical prote
81	28	70.0	178	2	D81269	hypothetical prote
82	28	70.0	183	2	T30677	hypothetical prote
83	28	70.0	192	2	S38879	LIM-domain protein
84	28	70.0	192	2	A49648	cysteine-rich prot
85	28	70.0	193	2	S53580	cysteine-rich prot
86	28	70.0	193	2	S12658	cysteine-rich prot
87	28	70.0	272	2	T24709	hypothetical prote
88	28	70.0	283	2	T32480	hypothetical prote
89	28	70.0	291	2	AB1397	hypothetical prote
90	28	70.0	296	2	A36966	probable dTDP-rham
91	28	70.0	302	2	AI1160	CheA activity-modu
92	28	70.0	302	2	AI1519	CheA activity-modu
93	28	70.0	332	2	T45723	hypothetical prote
94	28	70.0	334	2	T20562	hypothetical prote
95	28	70.0	339	2	G82596	phage-related prot
96	28	70.0	358	2	AE2425	hypothetical prote
97	28	70.0	420	2	A96535	unknown protein, 1
98	28	70.0	443	2	G81271	biotin carboxylase
99	28	70.0	466	2	T02324	hypothetical prote
100	28	70.0	470	2	AI3442	glycolate oxidase
101	28	70.0	486	2	T50749	methoxyneurosporen
102	28	70.0	486	2	S49624	methoxyneurosporen
103	28	70.0	495	2	S23633	methoxyneurosporen
104	28	70.0	501	2	E89786	hypothetical prote
105	28	70.0	540	2	T48619	tRNA synthase-like
106	28	70.0	541	2	S68249	chaperonin groEL h
107	28	70.0	551	2	E64728	yabN protein - Esc
108	28	70.0	552	2	B90638	probable transport
109	28	70.0	552	2	B85489	probable transport
110	28	70.0	559	2	G71327	probable apolipopr
111	28	70.0	575	2	JG0181	X11L2 protein - hu
112	28	70.0	592	2	T15413	hypothetical prote
113	28	70.0	624	1	BHTLE	hemocyanin chain e
114	28	70.0	631	1	BHTLA	hemocyanin chain a
115	28	70.0	728	2	S43768	transcription acti
116	28	70.0	754	2	T06249	protoporphyrin IX
117	28	70.0	758	2	T02925	protoporphyrin IX
118	28	70.0	778	2	B86218	protein T27G7.20 [
119	28	70.0	837	2	T27467	hypothetical prote
120	28	70.0	1031	2	F83561	probable DNA polym
121	28	70.0	1375	2	T18961	FAB1 protein homol
122	28	70.0	2643	2	T29149	hypothetical prote
123	28	70.0	3898	1	GNWVHB	genome polyprotein
124	28	70.0	3898	2	S57437	genome polyprotein
125	28	70.0	3898	2	S58295	polyprotein - hog
126	27	67.5	50	2	F95032	hypothetical prote
127	27	67.5	56	2	T07227	hypothetical prote
128	27	67.5	92	2	D83138	hypothetical prote

129	27	67.5	104	2	D72337	conserved hypothet
130	27	67.5	109	1	MNIHB2	nonstructural prot
131	27	67.5	109	1	A44275	nonstructural prot
132	27	67.5	109	2	S58182	nonstructural prot
133	27	67.5	109	2	S58186	nonstructural prot
134	27	67.5	127	2	H72566	probable ribosomal
135	27	67.5	142	2	T23519	hypothetical prote
136	27	67.5	147	2	T24057	hypothetical prote
137	27	67.5	149	2	T08218	hypothetical prote
138	27	67.5	153	2	D89940	hypothetical prote
139	27	67.5	154	2	S30728	hypothetical prote
140	27	67.5	163	2	F86878	hypothetical prote
141	27	67.5	167	2	F81263	probable lipoprote
142	27	67.5	174	2	AC1587	hypothetical prote
143	27	67.5	177	2	T16280	hypothetical prote
144	27	67.5	186	2	AE1622	B. subtilis ComEB
145	27	67.5	186	2	AC1260	B. subtilis ComEB
146	27	67.5	190	2	G84182	hypothetical prote
147	27	67.5	204	2	H70008	hypothetical prote
148	27	67.5	212	2	T47947	hypothetical prote
149	27	67.5	216	2	T12812	hypothetical prote
150	27	67.5	230	2	H81401	hydrophobic protei
151	27	67.5	235	2	T08467	aspartate racemase
152	27	67.5	236	2	A81690	conserved hypothet
153	27	67.5	248	2	AB0177	probable short cha
154	27	67.5	258	2	H65188	sec-independent pr
155	27	67.5	258	2	H91224	Sec-independent pr
156	27	67.5	258	2	F86071	sec-independent pr
157	27	67.5	271	2	B64932	hypothetical prote
158	27	67.5	273	2	AH2432	hypothetical prote
159	27	67.5	281	2	F69400	2-deoxy-D-gluconat
160	27	67.5	286	2	C84857	hypothetical prote
161	27	67.5	289	2	E86624	hypothetical prote
162	27	67.5	289	2	B72001	conserved hypothet
163	27	67.5	293	2	T10523	xyloglucan endo-1,
164	27	67.5	293	2	T48975	xyloglucan endo-tr
165	27	67.5	293	2	A71946	hypothetical prote
166	27	67.5	300	2	T26245	hypothetical prote
167	27	67.5	305	2	T50474	GTP-binding regula
168	27	67.5	310	2	S52504	probable membrane
169	27	67.5	313	1	H64596	chemotaxis protein
170	27	67.5	313	2	B71916	probable chemotaxi
171	27	67.5	314	2	E84046	hypothetical prote
172	27	67.5	317	2	S72457	GTP-binding regula
173	27	67.5	323	2	AC3514	hypothetical cytos
174	27	67.5	324	2	F97729	hypothetical prote
175	27	67.5	337	2	AG0963	hypothetical prote
176	27	67.5	338	2	JC7769	38K protein - huma
177	27	67.5	345	2	F95088	conserved hypothet
178	27	67.5	345	2	A97956	DNA polymerase III
179	27	67.5	346	2	AC3108	hypothetical prote
180	27	67.5	349	2	T26247	hypothetical prote
181	27	67.5	361	2	F96034	probable sugar upt
182	27	67.5	371	2	A98179	ABC transporter, A
183	27	67.5	374	2	AB1765	B. subtilis O-succ
184	27	67.5	385	2	G87340	aminotransferase,
185	27	67.5	387	2	S64082	probable membrane

186	27	67.5	387	2	S07749	hypothetical prote
187	27	67.5	394	2	S26431	intermediate filam
188	27	67.5	422	2	C95053	IS1167, transposas
189	27	67.5	432	2	S26432	intermediate filam
190	27	67.5	433	2	T08938	hypothetical prote
191	27	67.5	434	2	T26275	hypothetical prote
192	27	67.5	437	2	H64251	replication initia
193	27	67.5	445	2	T39350	glutamate n-acetyl
194	27	67.5	445	2	JC7861	caspase-associated
195	27	67.5	455	2	D69078	probable phosphoma
196	27	67.5	483	2	S69894	major DNA-binding
197	27	67.5	484	2	T22230	hypothetical prote
198	27	67.5	490	2	A96556	probable tRNA-guan
199	27	67.5	492	2	T38211	hypothetical prote
200	27	67.5	500	2	H91283	hypothetical prote
201	27	67.5	500	2	B86125	hypothetical prote
202	27	67.5	500	2	S56489	hypothetical 54.3K
203	27	67.5	513	2	AE0021	conserved hypothet
204	27	67.5	513	2	E71683	NADH2 dehydrogenas
205	27	67.5	516	2	A87260	conserved hypothet
206	27	67.5	521	2	AC1525	internalin like pr
207	27	67.5	522	2	H86248	protein T23J18.22
208	27	67.5	533	2	F70551	hypothetical prote
209	27	67.5	576	1	B70558	probable ABC trans
210	27	67.5	580	2	T40939	probable Ca-calmod
211	27	67.5	597	2	T01808	hypothetical prote
212	27	67.5	598	2	T05130	hypothetical prote
213	27	67.5	610	2	S59394	protein kinase RCK
214	27	67.5	631	1	JC4602	protein kinase (EC
215	27	67.5	631	2	E83778	serine protein kin
216	27	67.5	703	2	T04191	hypothetical prote
217	27	67.5	741	2	T46488	hypothetical prote
218	27	67.5	779	2	E71825	probable type III
219	27	67.5	783	2	A87438	DNA ligase, NAD-de
220	27	67.5	847	2	AG1949	hypothetical prote
221	27	67.5	879	2	E69792	conserved hypothet
222	27	67.5	904	2	T03806	hypothetical prote
223	27	67.5	930	2	A95193	isoleucyl-tRNA syn
224	27	67.5	930	2	E98059	isoleucine-tRNA li
225	27	67.5	949	2	D82293	isoleucyl-tRNA syn
226	27	67.5	955	2	F72243	hypothetical prote
227	27	67.5	995	2	T39724	probable pre-mRNA
228	27	67.5	996	2	E98092	cylM protein, cyto
229	27	67.5	996	2	A95228	bacteriocin format
230	27	67.5	1297	2	T39287	hypothetical prote
231	27	67.5	1451	2	T33083	hypothetical prote
232	27	67.5	1525	2	T01661	probable DNA (cyto
233	27	67.5	1647	2	T32934	hypothetical prote
234	27	67.5	1697	2	T00079	hypothetical prote
235	27	67.5	1713	2	A55347	adhesive ligand ep
236	27	67.5	2161	1	A45389	genome polyprotein
237	27	67.5	3068	1	A44062	genome polyprotein
238	27	67.5	3137	2	A37797	collagen alpha 3(V
239	27	67.5	3951	1	VFIHB1	F1 protein - avian
240	26	65.0	32	2	D64571	hypothetical prote
241	26	65.0	32	2	S67962	valine-tRNA ligase
242	26	65.0	80	2	G84773	hypothetical prote

243	26	65.0	85	2	AG3408	molybdopterin (mpt
244	26	65.0	100	2	AF1550	B. subtilis YneR p
245	26	65.0	104	2	AE2131	hypothetical prote
246	26	65.0	109	2	C64402	hypothetical prote
247	26	65.0	115	2	AE0523	conserved hypothet
248	26	65.0	116	2	E86633	hypothetical prote
249	26	65.0	120	2	E84271	30S ribosomal prot
250	26	65.0	122	2	E84154	cadmium-binding pr
251	26	65.0	132	2	T35965	conserved hypothet
252	26	65.0	137	2	F81342	transthyretin-like
253	26	65.0	140	2	B95049	hypothetical prote
254	26	65.0	140	2	H97919	(3R)-hydroxymyrist
255	26	65.0	144	2	D72219	conserved hypothet
256	26	65.0	144	2	A86722	hypothetical prote
257	26	65.0	156	2	F85495	hypothetical prote
258	26	65.0	156	2	B64735	yacC protein - Esc
259	26	65.0	156	2	F90644	hypothetical prote
260	26	65.0	156	2	AB2980	hypothetical prote
261	26	65.0	158	2	E86191	hypothetical prote
262	26	65.0	159	2	T32098	hypothetical prote
263	26	65.0	162	2	T17047	NADH2 dehydrogenas
264	26	65.0	162	2	G72015	hypothetical prote
265	26	65.0	162	2	F86608	hypothetical prote
266	26	65.0	175	2	T29180	hypothetical prote
267	26	65.0	182	2	T06198	lipoxygenase (EC 1
268	26	65.0	183	1	F69768	ribosomal-protein-
269	26	65.0	186	1	WMRZ19	19K globulin precu
270	26	65.0	186	2	JC4784	alpha-globulin pre
271	26	65.0	187	2	D64320	hypothetical prote
272	26	65.0	190	2	S48101	xyloglucan endo-1,
273	26	65.0	192	2	S74627	hypothetical prote
274	26	65.0	205	2	G72318	chemotaxis protein
275	26	65.0	208	2	G02090	cysteine-rich prot
276	26	65.0	208	2	S38745	cysteine-rich prot
277	26	65.0	214	2	A86670	hypothetical prote
278	26	65.0	214	2	T47892	hypothetical prote
279	26	65.0	215	2	T14779	hypothetical prote
280	26	65.0	221	2	C71056	hypothetical prote
281	26	65.0	221	2	D95277	hypothetical prote
282	26	65.0	225	2	S50458	hypothetical prote
283	26	65.0	226	2	T20447	hypothetical prote
284	26	65.0	234	2	E72252	translation initia
285	26	65.0	237	2	AC1953	rubrerythrin [impo
286	26	65.0	240	2	C71277	hypothetical prote
287	26	65.0	244	2	T16110	hypothetical prote
288	26	65.0	247	2	T32514	hypothetical prote
289	26	65.0	248	2	D71672	hypothetical prote
290	26	65.0	248	2	T20027	hypothetical prote
291	26	65.0	250	2	AF2881	lipase esterase [i
292	26	65.0	250	2	F97657	lipase esterase (A
293	26	65.0	251	2	C75200	hypothetical prote
294	26	65.0	252	2	T13588	NADH2 dehydrogenas
295	26	65.0	255	2	H71154	probable competenc
296	26	65.0	256	2	T00165	repressor - Staphy
297	26	65.0	257	2	S00682	IgE Fc receptor al
298	26	65.0	257	2	C69862	conserved hypothet
299	26	65.0	259	2	AG0915	sec-independent pr



300	26	65.0	259	2	T20205	hypothetical prote
301	26	65.0	260	2	G69900	hypothetical prote
302	26	65.0	260	2	G87349	conserved hypothet
303	26	65.0	264	2	B71448	hypothetical prote
304	26	65.0	265	2	T44414	hypothetical prote
305	26	65.0	267	2	T46202	endoxyloglucan tra
306	26	65.0	274	2	T26658	hypothetical prote
307	26	65.0	276	2	D87323	chemotaxis protein
308	26	65.0	278	2	B46290	mu-crystallin - hu
309	26	65.0	285	2	T01119	hypothetical prote
310	26	65.0	286	2	D90609	conserved hypothet
311	26	65.0	291	2	T34043	hypothetical prote
312	26	65.0	292	2	G72530	probable nicotine
313	26	65.0	295	2	S48102	xyloglucan endo-1,
314	26	65.0	299	1	ZBBE13	33.1K zinc-binding
315	26	65.0	299	2	B84941	ATP phosphoribosyl
316	26	65.0	300	2	E87428	conserved hypothet
317	26	65.0	300	2	T28014	hypothetical prote
318	26	65.0	303	2	F84220	citrate (pro-3S)-1
319	26	65.0	305	1	RNVZ35	DNA-directed RNA p
320	26	65.0	305	1	RNVZC5	DNA-directed RNA p
321	26	65.0	305	2	F72167	A33L protein - var
322	26	65.0	305	2	JQ1835	4L protein - vario
323	26	65.0	305	2	F36851	DNA-directed RNA p
324	26	65.0	305	2	T37415	RNA polymerase sub
325	26	65.0	305	2	T22009	hypothetical prote
326	26	65.0	306	2	A54717	palmitoyl-protein
327	26	65.0	314	1	QXBP2L	hypothetical prote
328	26	65.0	321	2	E81725	MesJ/Ycf62 family
329	26	65.0	324	2	JG0163	glucuronyltransfer
330	26	65.0	325	2	T24737	hypothetical prote
331	26	65.0	326	2	B84383	porphobilinogen sy
332	26	65.0	327	2	S26647	phosphoprotein - P
333	26	65.0	328	2	T36200	probable DNA polym
334	26	65.0	332	2	T25676	hypothetical prote
335	26	65.0	337	2	AB3603	3-methyl-2-oxobuta
336	26	65.0	338	2	S47972	dioxygenase, iron
337	26	65.0	338	2	E97086	anaerobic sulfite
338	26	65.0	340	1	MMBE5	cell fusion protei
339	26	65.0	342	2	E70109	hypothetical prote
340	26	65.0	344	2	E90564	potassium channel
341	26	65.0	347	2	JC7828	glucuronyltransfer
342	26	65.0	348	2	H89007	protein F59B1.1 [i
343	26	65.0	350	2	G75212	hypothetical prote
344	26	65.0	352	2	H97272	histidinol-phospha
345	26	65.0	352	2	AB3079	transcription regu
346	26	65.0	352	2	G98207	hypothetical prote
347	26	65.0	356	2	D84280	hypothetical prote
348	26	65.0	357	2	S75666	3-chlorobenzoate-3
349	26	65.0	361	2	G64354	hypothetical prote
350	26	65.0	365	2	C64309	hypothetical prote
351	26	65.0	368	2	AH3202	conserved hypothet
352	26	65.0	371	2	D83650	DNA repair and gen
353	26	65.0	381	2	T34333	hypothetical prote
354	26	65.0	386	2	A87697	DegT/DnrJ/EryC1/St
355	26	65.0	393	2	T13501	NADH2 dehydrogenas
356	26	65.0	397	1	Z6BPT9	ribonucleoside-tri

357	26	65.0	397	2	D89768	hypothetical prote
358	26	65.0	398	2	F72335	hypothetical prote
359	26	65.0	398	2	T44331	hypothetical prote
360	26	65.0	400	2	D64462	hypothetical prote
361	26	65.0	403	2	H70324	hypothetical prote
362	26	65.0	406	2	A43850	glutamate N-acetyl
363	26	65.0	406	2	C81017	amino-acid N-acety
364	26	65.0	410	2	E96707	hypothetical prote
365	26	65.0	417	2	S28922	phosphoglycerate k
366	26	65.0	419	2	A70156	RNA polymerase sig
367	26	65.0	420	2	S36444	hygromycin phospho
368	26	65.0	433	2	T17654	hypothetical prote
369	26	65.0	442	2	H81402	probable integral
370	26	65.0	446	2	S59646	clathrin coat asse
371	26	65.0	447	2	F71039	hypothetical prote
372	26	65.0	455	2	T34366	hypothetical prote
373	26	65.0	456	2	D95384	protein [imported
374	26	65.0	467	2	AE1521	hypothetical prote
375	26	65.0	468	2	S75389	probable phenylala
376	26	65.0	471	1	A41706	tryptophan-tRNA li
377	26	65.0	475	1	KCRTIH	stromelysin 1 (EC
378	26	65.0	475	1	YWBO	tryptophan-tRNA li
379	26	65.0	475	1	YWRBPR	tryptophan-tRNA li
380	26	65.0	476	2	G70175	aminoacyl-histidin
381	26	65.0	477	1	KCHUS1	stromelysin 1 (EC
382	26	65.0	477	1	KCMSS1	stromelysin 1 (EC
383	26	65.0	478	1	KCRBS1	stromelysin 1 (EC
384	26	65.0	478	2	A86142	protein T25K16.17
385	26	65.0	481	2	S50053	tryptophan-tRNA li
386	26	65.0	481	2	T20309	hypothetical prote
387	26	65.0	490	2	A35312	potassium channel
388	26	65.0	493	2	G69401	lysyl-tRNA synthet
389	26	65.0	501	2	S59946	hydroxymethylgluta
390	26	65.0	508	2	T07971	aromatic-L-amino-a
391	26	65.0	511	1	S55198	sulfate adenyllyltr
392	26	65.0	511	1	VGVN	spike glycoprotein
393	26	65.0	512	2	C91268	transcription acti
394	26	65.0	512	2	A86109	transcription acti
395	26	65.0	512	2	C41968	transcription acti
396	26	65.0	513	2	S26704	NADH2 dehydrogenas
397	26	65.0	518	2	S55948	hypothetical prote
398	26	65.0	531	2	T04463	hypothetical prote
399	26	65.0	543	2	G95118	RNA methyltransfer
400	26	65.0	543	2	D97988	conserved hypothet
401	26	65.0	544	2	T40058	probable chromatin
402	26	65.0	545	2	F64579	site-specific DNA-
403	26	65.0	545	2	F71932	type II DNA modifi
404	26	65.0	558	2	T47851	hypothetical prote
405	26	65.0	563	2	AD0488	probable membrane
406	26	65.0	572	2	H96685	probable AMP-bindi
407	26	65.0	574	2	T07112	hydroxymethylgluta
408	26	65.0	574	2	AB1790	ABC transporter re
409	26	65.0	574	2	AC1414	ABC transporter re
410	26	65.0	574	2	D91159	probable transport
411	26	65.0	574	2	B86005	probable transport
412	26	65.0	575	2	JC7727	4-alpha-D-((1->4)-
413	26	65.0	577	2	T50341	hypothetical prote

414	26	65.0	582	2	I38028	matrix metalloprot
415	26	65.0	582	2	G81288	hypothetical prote
416	26	65.0	595	2	D83806	aspartyl-tRNA synt
417	26	65.0	596	2	S59944	hydroxymethylgluta
418	26	65.0	601	2	H69274	probable long-chai
419	26	65.0	602	1	S25316	hydroxymethylgluta
420	26	65.0	604	2	T04120	hydroxymethylgluta
421	26	65.0	604	2	S24760	hydroxymethylgluta
422	26	65.0	613	2	B96501	hypothetical prote
423	26	65.0	615	2	C97723	aspartate-tRNA lig
424	26	65.0	620	2	H69382	ABC transporter, A
425	26	65.0	626	2	S67964	hemocyanin chain 6
426	26	65.0	637	2	H85043	hypothetical prote
427	26	65.0	643	2	E82481	methyl-accepting c
428	26	65.0	660	2	T04416	sulfate transport
429	26	65.0	664	2	D81330	glycine-tRNA ligas
430	26	65.0	670	2	T33304	hypothetical prote
431	26	65.0	692	2	C87644	fatty oxidation co
432	26	65.0	704	2	A34287	storage protein 2
433	26	65.0	705	2	B82885	hypothetical prote
434	26	65.0	708	2	T39742	hypothetical prote
435	26	65.0	716	2	B72258	hypothetical prote
436	26	65.0	721	2	T46037	hypothetical prote
437	26	65.0	726	2	T15810	hypothetical prote
438	26	65.0	733	2	D83588	conserved hypothet
439	26	65.0	735	2	D70174	methyl-accepting c
440	26	65.0	749	2	A45046	basic juvenile hor
441	26	65.0	762	2	H87466	beta-D-glucosidase
442	26	65.0	768	2	B41029	integrin beta-8 ch
443	26	65.0	769	2	A41029	integrin beta-8 ch
444	26	65.0	776	2	B96666	protein F22C12.6 [
445	26	65.0	788	2	C84616	similar to mammali
446	26	65.0	796	2	T03746	hypothetical prote
447	26	65.0	804	2	T25673	hypothetical prote
448	26	65.0	811	2	S39901	nwsA protein - Bra
449	26	65.0	855	2	S47533	glucose-6-phosphat
450	26	65.0	858	2	D83491	probable sensor/re
451	26	65.0	862	2	T05941	lipoxygenase (EC 1
452	26	65.0	864	2	T05945	lipoxygenase (EC 1
453	26	65.0	865	2	A25762	regulatory protein
454	26	65.0	873	2	T05189	glutamyl aminopept
455	26	65.0	875	2	T33434	hypothetical prote
456	26	65.0	888	2	T01081	hypothetical prote
457	26	65.0	898	2	S76431	endopeptidase Clp
458	26	65.0	904	1	RGBYP1	regulatory protein
459	26	65.0	904	2	G90563	lipoprotein [impor
460	26	65.0	906	2	T48898	disease resistance
461	26	65.0	908	2	T48899	disease resistance
462	26	65.0	910	2	S40259	glucose-6-phosphat
463	26	65.0	919	2	S33942	hexon protein - hu
464	26	65.0	947	2	T02979	pyruvate, phosphat
465	26	65.0	964	2	T15342	hypothetical prote
466	26	65.0	969	2	S17909	hypothetical prote
467	26	65.0	977	2	AC1963	DNA polymerase I [
468	26	65.0	980	2	S54986	regulatory protein
469	26	65.0	990	2	T03784	probable receptor
470	26	65.0	1016	2	JQ0498	genome polyprotein

471	26	65.0	1020	2	G88208	protein K02A2.3 [i
472	26	65.0	1045	2	B30239	hydroxymethylgluta
473	26	65.0	1078	2	T30879	dynein heavy chain
474	26	65.0	1088	2	T41671	hypothetical prote
475	26	65.0	1119	2	A86340	protein F2D10.24 [
476	26	65.0	1121	2	T21303	hypothetical prote
477	26	65.0	1144	1	A43271	nitric-oxide synth
478	26	65.0	1163	2	D64315	type I restriction
479	26	65.0	1174	2	H84982	exodeoxyribonuclea
480	26	65.0	1175	2	D85089	hypothetical prote
481	26	65.0	1220	2	T32916	hypothetical prote
482	26	65.0	1251	2	S49645	probable membrane
483	26	65.0	1280	2	T51500	hypothetical prote
484	26	65.0	1285	2	B72420	hypothetical prote
485	26	65.0	1318	2	T39066	hypothetical prote
486	26	65.0	1325	1	S73723	probable lipoprote
487	26	65.0	1365	2	S14871	suppressor two of
488	26	65.0	1490	2	F88311	protein T06D8.10 [
489	26	65.0	1490	2	T24502	hypothetical prote
490	26	65.0	1519	2	S41525	major ring-forming
491	26	65.0	2140	2	F95074	serine proteinase,
492	26	65.0	2144	2	A97942	metalloproteinase
493	26	65.0	2199	2	T40008	Cdc20p - fission y
494	26	65.0	2366	2	S10317	toxin B - Clostrid
495	26	65.0	2367	2	S70172	toxin B - Clostrid
496	26	65.0	2670	2	A46719	inositol 1,4,5-tri
497	26	65.0	2671	2	A49873	inositol 1,4,5-tri
498	26	65.0	3066	1	JQ1661	genome polyprotein
499	26	65.0	3066	1	JQ1662	genome polyprotein
500	25	62.5	37	2	F83931	hypothetical prote
501	25	62.5	49	2	C82788	hypothetical prote
502	25	62.5	69	2	C72262	hypothetical prote
503	25	62.5	77	2	A99925	hypothetical prote
504	25	62.5	81	2	H64814	molybdopterin bios
505	25	62.5	81	2	A82251	molybdenum cofacto
506	25	62.5	81	2	F90736	molybdopterin bios
507	25	62.5	81	2	G85586	molybdopterin bios
508	25	62.5	86	2	H82855	hypothetical prote
509	25	62.5	88	2	I47758	retrovirus-related
510	25	62.5	98	2	S54735	ribosomal protein
511	25	62.5	99	2	AH1602	conserved hypothet
512	25	62.5	104	2	F72590	hypothetical prote
513	25	62.5	110	2	T12461	hypothetical prote
514	25	62.5	111	2	T12857	hypothetical prote
515	25	62.5	111	2	S64475	hypothetical prote
516	25	62.5	117	2	S07933	hypothetical prote
517	25	62.5	117	2	B81181	hypothetical prote
518	25	62.5	117	2	A81925	hypothetical prote
519	25	62.5	126	2	H69267	cytochrome-c oxida
520	25	62.5	126	2	E96547	hypothetical prote
521	25	62.5	127	2	E97294	uncharacterized pr
522	25	62.5	131	2	F82331	conserved hypothet
523	25	62.5	132	2	I40566	hypothetical prote
524	25	62.5	132	2	A97204	probable aldoketom
525	25	62.5	140	2	G84116	hydroxymyristoyl-(
526	25	62.5	142	2	F86893	competence regulat
527	25	62.5	149	2	F97485	hypothetical prote

528	25	62.5	150	2	T29939	hypothetical prote
529	25	62.5	153	1	QQVZF7	F7 protein - vacci
530	25	62.5	153	2	D42513	J1R protein - vacc
531	25	62.5	153	2	T37361	dimeric virion pro
532	25	62.5	156	2	T33625	hypothetical prote
533	25	62.5	157	2	S76232	hypothetical prote
534	25	62.5	157	2	B83066	hypothetical prote
535	25	62.5	159	2	D72160	M1R protein - vari
536	25	62.5	159	2	S33092	J1R protein - vari
537	25	62.5	159	2	T28516	hypothetical prote
538	25	62.5	159	2	I55083	2C-methyl-D-erythr
539	25	62.5	159	2	A85924	2C-methyl-D-erythr
540	25	62.5	159	2	AD0856	2C-methyl-D-erythr
541	25	62.5	159	2	H91078	2C-methyl-D-erythr
542	25	62.5	164	2	T17050	NADH2 dehydrogenas
543	25	62.5	164	2	S29392	estradiol-stimulat
544	25	62.5	165	2	S15800	hypothetical prote
545	25	62.5	170	2	G81375	hypothetical prote
546	25	62.5	172	2	F88542	protein ZK637.15 [
547	25	62.5	176	2	B44457	RNA polymerase II
548	25	62.5	176	2	A44457	DNA-directed RNA p
549	25	62.5	176	2	A97808	hypothetical prote
550	25	62.5	177	2	H97195	serine/threonine k
551	25	62.5	178	2	G84357	probable acetyltra
552	25	62.5	179	2	H64472	hypothetical prote
553	25	62.5	181	2	E69900	hypothetical prote
554	25	62.5	183	2	T28929	hypothetical prote
555	25	62.5	185	2	S76844	hypothetical prote
556	25	62.5	188	2	S77089	hypothetical prote
557	25	62.5	190	2	A81334	recombination prot
558	25	62.5	190	2	B97827	hypothetical prote
559	25	62.5	192	2	I80320	hypothetical 22K p
560	25	62.5	193	2	F97212	probable phosphata
561	25	62.5	194	2	F64025	hypothetical prote
562	25	62.5	196	2	I39698	blue copper-bindin
563	25	62.5	196	2	T51838	blue copper bindin
564	25	62.5	196	2	T07458	endopeptidase Clp
565	25	62.5	196	2	D70637	hypothetical prote
566	25	62.5	197	2	T29930	hypothetical prote
567	25	62.5	197	2	AI2184	hypothetical prote
568	25	62.5	202	2	G86520	glucose-6-P dehydro
569	25	62.5	203	2	A96595	hypothetical prote
570	25	62.5	205	1	S50763	endopeptidase Clp
571	25	62.5	206	2	A86869	transport protein
572	25	62.5	210	2	G97037	hypothetical prote
573	25	62.5	213	2	H88542	protein ZK637.12 [
574	25	62.5	217	2	D97321	response regulator
575	25	62.5	219	2	AI2680	conserved hypothet
576	25	62.5	224	2	G71483	hypothetical prote
577	25	62.5	224	2	E86321	hypothetical prote
578	25	62.5	228	2	T47847	hypothetical prote
579	25	62.5	228	2	T48549	hypothetical prote
580	25	62.5	230	2	F64505	phosphoribosylform
581	25	62.5	230	2	B82195	hypothetical prote
582	25	62.5	233	2	T10198	hypothetical prote
583	25	62.5	235	2	S39736	ywfC protein - Bac
584	25	62.5	237	2	F70446	conserved hypothet

585	25	62.5	242	2	T27226	hypothetical prote
586	25	62.5	243	2	A70670	hypothetical prote
587	25	62.5	245	2	AG1461	probable phospho-b
588	25	62.5	245	2	AH1098	a probable phospho
589	25	62.5	247	2	T37001	hypothetical prote
590	25	62.5	251	2	F96008	hypothetical prote
591	25	62.5	253	1	F71233	hypothetical prote
592	25	62.5	253	2	G97462	hypothetical prote
593	25	62.5	256	2	AC1561	conserved hypothet
594	25	62.5	256	2	AE1203	conserved hypothet
595	25	62.5	258	2	D72103	glucose-6-p dehydro
596	25	62.5	259	2	T43102	hypothetical prote
597	25	62.5	260	2	B84012	N-acetylglutamate
598	25	62.5	263	2	B70153	conserved hypothet
599	25	62.5	266	2	T29609	hypothetical prote
600	25	62.5	269	2	T22778	hypothetical prote
601	25	62.5	271	1	A43744	N-acylmannosamine
602	25	62.5	273	2	E82655	conserved hypothet
603	25	62.5	274	2	H83707	bacitracin resista
604	25	62.5	278	2	A83840	undecaprenol kinas
605	25	62.5	279	2	D70152	hydrolase homolog
606	25	62.5	280	2	A81746	chromosome partion
607	25	62.5	280	2	AI2704	conserved hypothet
608	25	62.5	281	2	G70309	protoporphyrinogen
609	25	62.5	285	2	A97487	probable iron-sulf
610	25	62.5	289	2	G96981	ABC-type sugar tra
611	25	62.5	290	2	T24926	hypothetical prote
612	25	62.5	290	2	E90389	hypothetical prote
613	25	62.5	291	2	AD1772	hypothetical prote
614	25	62.5	292	2	F81880	probable 5,10-meth
615	25	62.5	292	2	D81140	5,10-methylenetetra
616	25	62.5	297	2	G69525	formylmethanofuran
617	25	62.5	300	2	T38986	probable c-4 methy
618	25	62.5	301	2	S39679	transcription regu
619	25	62.5	301	2	G71206	tryptophan-tRNA li
620	25	62.5	303	2	T23583	hypothetical prote
621	25	62.5	303	2	B96909	probable permease
622	25	62.5	305	2	T33578	hypothetical prote
623	25	62.5	306	2	D86866	prophage ps3 prote
624	25	62.5	309	2	A71113	hypothetical prote
625	25	62.5	309	2	AD3594	transporter, dme f
626	25	62.5	316	2	C70438	conserved hypothet
627	25	62.5	317	2	H82785	dolichol-phosphate
628	25	62.5	318	2	E87637	hypothetical prote
629	25	62.5	318	2	C85071	hypothetical prote
630	25	62.5	319	2	AG1169	hypothetical prote
631	25	62.5	319	2	AI1526	hypothetical prote
632	25	62.5	321	2	E97741	D-alanine-D-alanin
633	25	62.5	321	2	D69101	coenzyme F420-depe
634	25	62.5	321	2	S66529	N5,N10-methylenete
635	25	62.5	323	2	H86716	L-asparaginase [im
636	25	62.5	324	2	H96746	RING-H2 zinc finge
637	25	62.5	331	2	T26871	hypothetical prote
638	25	62.5	332	2	C83423	conserved hypothet
639	25	62.5	334	2	B72301	endoglucanase - Th
640	25	62.5	334	2	C84964	hypothetical prote
641	25	62.5	336	2	AC1979	ferric iron-bindin

642	25	62.5	336	2	G84025	polysugar degradin
643	25	62.5	337	2	C72399	DNA processing cha
644	25	62.5	338	2	T05036	xyloglucan endo-1,
645	25	62.5	338	2	B95290	hypothetical prote
646	25	62.5	340	2	B84019	hypothetical prote
647	25	62.5	340	2	H81346	hypothetical prote
648	25	62.5	341	2	S71223	xyloglucan endo-1,
649	25	62.5	342	2	E69581	acetoin dehydrogen
650	25	62.5	343	1	C70418	probable alcohol d
651	25	62.5	343	2	B86446	probable endoxylog
652	25	62.5	346	2	G96643	hypothetical prote
653	25	62.5	346	2	T31719	hypothetical prote
654	25	62.5	346	2	D81288	probable fucose sy
655	25	62.5	347	2	T20618	hypothetical prote
656	25	62.5	350	2	B39364	GDF-1 embryonic gr
657	25	62.5	352	2	E70847	hypothetical prote
658	25	62.5	354	2	F90121	hypothetical prote
659	25	62.5	357	2	AF1729	protein gp20 (Bact
660	25	62.5	359	2	I51372	angiotensin II rec
661	25	62.5	359	2	C84983	hypothetical prote
662	25	62.5	362	2	S45887	ribosomal protein
663	25	62.5	362	2	S50993	ribosomal protein
664	25	62.5	364	1	C69351	probable iron-sulf
665	25	62.5	367	2	A84998	hypothetical prote
666	25	62.5	368	2	F70327	conserved hypothet
667	25	62.5	376	1	ATAXE	actin - Entamoeba
668	25	62.5	378	2	C83853	cystathionine gamm
669	25	62.5	379	2	S27502	hypothetical prote
670	25	62.5	380	2	F70399	hydrogenase expres
671	25	62.5	381	2	G64047	cystathionine gamm
672	25	62.5	381	2	E88427	protein R07E5.3 [i
673	25	62.5	383	2	T12656	NADH2 dehydrogenas
674	25	62.5	383	2	S76334	hypothetical prote
675	25	62.5	384	2	AE1636	N-acetylornithine
676	25	62.5	385	2	C75020	tryptophanyl-tRNA
677	25	62.5	386	2	G75079	hypothetical prote
678	25	62.5	386	2	B71066	hypothetical prote
679	25	62.5	387	2	T40704	hypothetical prote
680	25	62.5	388	2	T32307	hypothetical prote
681	25	62.5	388	2	B89777	capsular polysacch
682	25	62.5	389	2	F86212	hypothetical prote
683	25	62.5	390	2	C75103	na+/h+ antiporter
684	25	62.5	393	2	T13775	NADH2 dehydrogenas
685	25	62.5	393	2	S57671	hypothetical prote
686	25	62.5	395	2	T13756	NADH2 dehydrogenas
687	25	62.5	395	2	T32309	hypothetical prote
688	25	62.5	400	2	H69009	hypothetical prote
689	25	62.5	400	2	H87444	hypothetical prote
690	25	62.5	401	2	B70398	argininosuccinate
691	25	62.5	402	2	T13499	NADH2 dehydrogenas
692	25	62.5	404	2	G69332	heterodisulfide re
693	25	62.5	406	2	AC2267	hypothetical prote
694	25	62.5	408	2	C96984	probable Mn transp
695	25	62.5	417	2	S51961	FUN50 protein - ye
696	25	62.5	418	2	S31124	hypothetical prote
697	25	62.5	422	2	A71147	hypothetical prote
698	25	62.5	422	2	AC1189	phosphotransferase

699	25	62.5	429	2	D64499	glycine hydroxymet
700	25	62.5	429	2	T36088	probable secreted
701	25	62.5	430	2	G70487	sulfide-quinone re
702	25	62.5	431	2	B81254	probable transmemb
703	25	62.5	434	2	T04248	hypothetical prote
704	25	62.5	436	2	E96635	hypothetical prote
705	25	62.5	440	2	A71027	probable NADH oxid
706	25	62.5	443	2	G64234	hypothetical prote
707	25	62.5	443	2	C88427	protein R07E5.6 [i
708	25	62.5	443	2	T45574	anthranilate N-hyd
709	25	62.5	447	2	G84687	probable disease r
710	25	62.5	448	2	B69965	D-serine dehydrata
711	25	62.5	453	2	B70316	DAPA aminotransfer
712	25	62.5	453	2	F75206	maltose-binding pe
713	25	62.5	457	2	AB1426	GTPase homolog lmo
714	25	62.5	457	2	AH1799	GTPase homolog lin
715	25	62.5	457	2	G90354	hypothetical prote
716	25	62.5	462	2	S74515	LlaI restriction a
717	25	62.5	463	2	E86464	hypothetical prote
718	25	62.5	464	2	C64462	adenosylmethionine
719	25	62.5	467	2	T27108	hypothetical prote
720	25	62.5	469	2	T17191	dihydrolipoamide d
721	25	62.5	471	2	T21102	hypothetical prote
722	25	62.5	472	1	S55379	cytochrome P450 CY
723	25	62.5	473	2	H83676	pyruvate dehydroge
724	25	62.5	473	2	B21159	cell surface antig
725	25	62.5	473	2	C84979	hypothetical prote
726	25	62.5	474	2	S50373	MEC3 protein - yea
727	25	62.5	475	1	RKITL	ribulose-bisphosph
728	25	62.5	477	2	T29592	hypothetical prote
729	25	62.5	478	2	F89651	protein T04F8.2 [i
730	25	62.5	479	2	F86285	F9L1.11 protein -
731	25	62.5	480	2	B64308	hypothetical prote
732	25	62.5	485	2	T10792	amidophosphoribosy
733	25	62.5	488	2	S78236	ribulose-bisphosph
734	25	62.5	490	2	F86841	iron-binding oxida
735	25	62.5	491	2	T48991	hypothetical prote
736	25	62.5	493	2	G90604	hypothetical prote
737	25	62.5	496	2	T20926	hypothetical prote
738	25	62.5	497	1	S20174	protein kinase MEK
739	25	62.5	497	2	S60161	transcription fact
740	25	62.5	497	2	T32090	hypothetical prote
741	25	62.5	499	2	JH0313	potassium channel
742	25	62.5	501	2	A21159	cell surface antig
743	25	62.5	501	2	S44258	sucrose-6-phosphat
744	25	62.5	502	2	T24471	hypothetical prote
745	25	62.5	502	2	T05135	hypothetical prote
746	25	62.5	503	1	VMUT1B	variant surface gl
747	25	62.5	503	2	AD3423	ATPase BMEI1370 [i
748	25	62.5	504	2	S48550	hypothetical prote
749	25	62.5	505	2	D69382	conserved hypothet
750	25	62.5	510	1	A35251	histidine ammonia-
751	25	62.5	510	2	E71695	hypothetical prote
752	25	62.5	512	2	S19439	probable membrane
753	25	62.5	518	2	A55066	tyrosine decarboxy
754	25	62.5	520	2	T48988	hypothetical prote
755	25	62.5	520	2	AF2370	serine/threonine k



756	25	62.5	523	2	T09615	tyrosine decarboxy
757	25	62.5	523	4	D55066	probable tyrosine
758	25	62.5	525	2	E96786	protein F10A5.13 [
759	25	62.5	525	2	T35084	hypothetical prote
760	25	62.5	525	2	T50893	methoxyneurosporen
761	25	62.5	527	2	A97763	lysine-tRNA ligase
762	25	62.5	528	2	D71694	lysine-tRNA ligase
763	25	62.5	529	2	S43599	Snf5 homolog R07E5
764	25	62.5	532	2	PN0108	RNA-directed RNA p
765	25	62.5	533	2	A84512	hypothetical prote
766	25	62.5	534	2	C64367	hypothetical prote
767	25	62.5	534	2	B87382	peptide chain rele
768	25	62.5	535	2	D96680	hypothetical prote
769	25	62.5	536	2	A71491	probable integral
770	25	62.5	538	2	T51734	neoxanthin cleavag
771	25	62.5	538	2	T49193	neoxanthin cleavag
772	25	62.5	546	2	S48932	hypothetical prote
773	25	62.5	548	2	A81650	conserved hypothet
774	25	62.5	554	2	A56730	carl protein - Pod
775	25	62.5	555	2	C83444	probable AMP-bindi
776	25	62.5	557	2	T04746	hypothetical prote
777	25	62.5	558	2	S61604	probable membrane
778	25	62.5	562	2	JU0033	hypothetical L1 pr
779	25	62.5	573	2	AB2401	hypothetical prote
780	25	62.5	575	2	S14955	hydroxymethylgluta
781	25	62.5	576	2	AF2361	flavoprotein [impo
782	25	62.5	586	1	RNEGB1	DNA-directed RNA p
783	25	62.5	592	2	S43597	coiled-coil protei
784	25	62.5	593	2	T19169	hypothetical prote
785	25	62.5	602	2	T37254	acetylcholinestera
786	25	62.5	604	2	F89453	protein F35H12.4 [
787	25	62.5	609	2	T11774	NADH2 dehydrogenas
788	25	62.5	610	2	T11544	NADH2 dehydrogenas
789	25	62.5	610	2	T23836	hypothetical prote
790	25	62.5	614	2	T40652	hypothetical prote
791	25	62.5	622	2	T24632	hypothetical prote
792	25	62.5	624	2	H90071	hypothetical prote
793	25	62.5	636	2	G97029	nH(3)-dependent NA
794	25	62.5	640	2	C72351	hypothetical prote
795	25	62.5	646	2	G84854	hypothetical prote
796	25	62.5	655	2	G96524	protein T1N15.9 [i
797	25	62.5	664	2	G86692	penicillin-binding
798	25	62.5	664	2	AG2960	glycosidase Atu328
799	25	62.5	664	2	G98322	hypothetical prote
800	25	62.5	680	2	T27078	hypothetical prote
801	25	62.5	693	2	H71062	hypothetical prote
802	25	62.5	696	2	D81346	probable iron upta
803	25	62.5	712	2	S23650	retrovirus-related
804	25	62.5	713	2	JC5870	poly(beta-D-mannur
805	25	62.5	717	2	T28829	hypothetical prote
806	25	62.5	736	2	G72621	probable translati
807	25	62.5	736	2	C69451	cationic amino aci
808	25	62.5	738	2	S59842	probable membrane
809	25	62.5	742	1	H75575	probable oxidoredu
810	25	62.5	747	2	I51579	complement factor
811	25	62.5	755	2	B95342	NosR Regulatory pr
812	25	62.5	770	2	PN0105	RNA-directed RNA p

813	25	62.5	780	2	T22555	hypothetical prote
814	25	62.5	781	2	JC7382	DNA-directed DNA p
815	25	62.5	787	2	G81692	inner membrane pro
816	25	62.5	789	2	T45762	vacuolar sorting p
817	25	62.5	790	2	T05576	hypothetical prote
818	25	62.5	791	2	S67265	hypothetical prote
819	25	62.5	792	2	G86564	phenylalanyl tRNA
820	25	62.5	792	2	D81608	phenylalanine-tRNA
821	25	62.5	792	2	H72058	phenylalanine-tRNA
822	25	62.5	792	2	T21276	hypothetical prote
823	25	62.5	803	2	E81804	hypothetical prote
824	25	62.5	804	2	G71546	probable DNA gyras
825	25	62.5	815	2	T08450	hypothetical prote
826	25	62.5	826	2	T28858	hypothetical prote
827	25	62.5	831	2	AB3513	ATPase virB4 homol
828	25	62.5	833	1	A31593	heat shock transcr
829	25	62.5	834	2	T06055	hypothetical prote
830	25	62.5	834	2	D86842	hypothetical prote
831	25	62.5	839	2	G96719	probable chromomet
832	25	62.5	840	2	S74707	nitrogen fixation
833	25	62.5	841	2	JC5894	killer cell inhibi
834	25	62.5	845	2	T07039	Hcr9-0 protein - t
835	25	62.5	853	2	T17461	disease resistance
836	25	62.5	855	2	T17460	disease resistance
837	25	62.5	858	2	JC2309	chitin synthase (E
838	25	62.5	860	2	T27084	hypothetical prote
839	25	62.5	878	2	T43767	DNA topoisomerase
840	25	62.5	881	2	T22738	hypothetical prote
841	25	62.5	884	1	IJMSCE	E-cadherin precurs
842	25	62.5	884	2	S34438	uvomorulin - mouse
843	25	62.5	884	2	AE3166	ATP-dependent DNA
844	25	62.5	902	2	T00588	hypothetical prote
845	25	62.5	921	2	S39295	beta-adaptin 1 - f
846	25	62.5	923	1	HXAD40	hexon protein - hu
847	25	62.5	927	2	T08034	serine/threonine p
848	25	62.5	947	1	KIZMPO	pyruvate, phosphat
849	25	62.5	952	1	HXAD5	hexon protein - hu
850	25	62.5	955	2	S56649	pyruvate, phosphat
851	25	62.5	956	1	S53297	pyruvate, phosphat
852	25	62.5	967	1	HXAD2	hexon protein - hu
853	25	62.5	972	2	T39876	hypothetical prote
854	25	62.5	973	2	T40778	hypothetical 129.5
855	25	62.5	991	2	E83137	probable nonriboso
856	25	62.5	1035	2	T16588	hypothetical prote
857	25	62.5	1067	2	T06312	hypothetical prote
858	25	62.5	1085	2	C96797	unknown protein [i
859	25	62.5	1088	2	B85068	UV-damaged DNA bin
860	25	62.5	1102	2	T12681	hypothetical prote
861	25	62.5	1155	2	D70148	DNA-directed RNA p
862	25	62.5	1181	2	B64516	hypothetical prote
863	25	62.5	1253	2	F86436	hypothetical prote
864	25	62.5	1258	2	T14853	reverse transcript
865	25	62.5	1258	2	T14855	reverse transcript
866	25	62.5	1259	4	GNHUL1	retrovirus-related
867	25	62.5	1260	4	GNLRL1	retrovirus-related
868	25	62.5	1275	2	B28096	line-1 protein ORF
869	25	62.5	1275	2	I38588	reverse transcript

870	25	62.5	1275	2	S65824	reverse transcript
871	25	62.5	1280	2	B34087	hypothetical prote
872	25	62.5	1304	2	T19397	hypothetical prote
873	25	62.5	1333	1	XOHUDH	xanthine dehydroge
874	25	62.5	1342	1	RNECB	DNA-directed RNA p
875	25	62.5	1342	1	S32680	DNA-directed RNA p
876	25	62.5	1342	1	RNEBBT	DNA-directed RNA p
877	25	62.5	1342	2	F91242	RNA polymerase bet
878	25	62.5	1342	2	A84934	DNA-directed RNA p
879	25	62.5	1342	2	C86090	RNA polymerase, be
880	25	62.5	1342	2	AC0456	DNA-directed RNA p
881	25	62.5	1342	2	AE0933	DNA-directed RNA p
882	25	62.5	1352	2	G71051	probable ATP-depen
883	25	62.5	1357	2	S61187	probable membrane
884	25	62.5	1389	2	T30824	RNA polymerase bet
885	25	62.5	1392	2	D82007	DNA-directed RNA p
886	25	62.5	1394	2	A81236	DNA-directed RNA p
887	25	62.5	1458	2	S36014	dynein heavy chain
888	25	62.5	1458	2	A45665	adult-specific bru
889	25	62.5	1466	2	T39557	vacuolar protein s
890	25	62.5	1564	2	T27121	hypothetical prote
891	25	62.5	1585	2	B69948	phage-related prot
892	25	62.5	1605	2	T31435	DNA-directed RNA p
893	25	62.5	1622	2	T45240	hypothetical prote
894	25	62.5	1785	2	S53976	probable membrane
895	25	62.5	1821	2	AG2335	hypothetical prote
896	25	62.5	1847	2	T28969	hypothetical prote
897	25	62.5	2021	2	A97859	190-KDa cell surfa
898	25	62.5	2149	2	C96695	ribulose biphosph
899	25	62.5	2241	2	T02857	conserved hypothet
900	25	62.5	2249	2	A41477	190K surface antig
901	25	62.5	2405	2	T08164	dynein alpha heavy
902	25	62.5	2535	2	T04824	hypothetical prote
903	25	62.5	2809	2	T30213	G-cadherin - sea u
904	25	62.5	2911	2	T20566	hypothetical prote
905	25	62.5	2958	2	S64921	probable membrane
906	25	62.5	4367	1	B54802	dynein heavy chain
907	25	62.5	4568	2	T08030	dynein beta heavy
908	25	62.5	4644	1	A38905	dynein heavy chain
909	25	62.5	4919	2	T31105	hypothetical prote
910	25	62.5	4936	2	AH2515	hypothetical prote
911	24.5	61.3	387	2	H71708	pol (A) POLYMERASE
912	24.5	61.3	436	2	G97701	polynucleotide ade
913	24.5	61.3	603	2	F72237	conserved hypothet
914	24	60.0	34	2	E64695	hypothetical prote
915	24	60.0	46	2	F82074	hypothetical prote
916	24	60.0	50	2	H82456	hypothetical prote
917	24	60.0	66	2	AE3608	hypothetical prote
918	24	60.0	66	2	AH2817	secretion protein
919	24	60.0	67	2	G95310	probable CspA6 col
920	24	60.0	67	2	AB3074	cold shock protein
921	24	60.0	71	2	A97596	sece (AF176664) [i
922	24	60.0	73	2	A38883	rab protein Rab6 -
923	24	60.0	74	2	E64214	hypothetical prote
924	24	60.0	75	2	E70158	hypothetical prote
925	24	60.0	76	2	G98212	probable cold shoc
926	24	60.0	77	1	GYRTI	cysteine-rich inte

927	24	60.0	77	2	JC2431	cysteine-rich prot
928	24	60.0	77	2	G02666	cysteine-rich prot
929	24	60.0	80	2	F96787	protein T4012.7 [i
930	24	60.0	81	2	F69291	hypothetical prote
931	24	60.0	84	2	T29254	hypothetical prote
932	24	60.0	84	2	F88451	protein K10D2.7 [i
933	24	60.0	86	1	R5MX23	ribosomal protein
934	24	60.0	86	1	C64322	ribosomal protein
935	24	60.0	90	2	D82760	hypothetical prote
936	24	60.0	94	2	D71284	hypothetical prote
937	24	60.0	95	2	AD0480	hypothetical prote
938	24	60.0	98	2	C71560	probable chltr t2
939	24	60.0	100	2	E36905	ribosomal protein
940	24	60.0	101	2	F86720	probable ribosomal
941	24	60.0	101	2	S45299	probable reverse t
942	24	60.0	106	2	E82605	conjugal transfer
943	24	60.0	106	2	I40737	T2 protein - Chlam
944	24	60.0	109	2	B72213	conserved hypothet
945	24	60.0	112	2	A69268	hypothetical prote
946	24	60.0	112	2	T13322	hypothetical prote
947	24	60.0	112	2	T47326	hypothetical prote
948	24	60.0	123	2	A72356	conserved hypothet
949	24	60.0	130	2	A69132	ribosomal protein
950	24	60.0	137	2	T12880	hypothetical prote
951	24	60.0	138	1	H64434	hypothetical prote
952	24	60.0	138	2	C97047	hypothetical prote
953	24	60.0	138	2	G97814	hypothetical prote
954	24	60.0	141	2	H69044	peptidylprolyl iso
955	24	60.0	143	2	B69203	conserved hypothet
956	24	60.0	144	2	G69425	hypothetical prote
957	24	60.0	145	2	AD1651	hypothetical prote
958	24	60.0	145	2	AG1588	hypothetical prote
959	24	60.0	145	2	AD1093	hypothetical prote
960	24	60.0	147	2	S65579	hypothetical prote
961	24	60.0	151	2	AI2892	hypothetical prote
962	24	60.0	153	2	T37219	probable lipoprote
963	24	60.0	156	2	H95252	PTS system, IIB co
964	24	60.0	156	2	F98117	hypothetical prote
965	24	60.0	156	2	T28141	C type lectin, B 1
966	24	60.0	158	2	A56125	placental growth f
967	24	60.0	158	2	AI2718	conserved hypothet
968	24	60.0	158	2	E97500	hypothetical prote
969	24	60.0	159	2	C72076	hypothetical prote
970	24	60.0	159	2	D86548	hypothetical prote
971	24	60.0	161	1	VCTMCP	coat protein - sun
972	24	60.0	162	2	T17044	NADH2 dehydrogenas
973	24	60.0	162	2	T17049	NADH2 dehydrogenas
974	24	60.0	162	2	T17045	NADH2 dehydrogenas
975	24	60.0	162	2	T14256	NADH2 dehydrogenas
976	24	60.0	162	2	T14247	NADH2 dehydrogenas
977	24	60.0	162	2	T17041	NADH2 dehydrogenas
978	24	60.0	162	2	T14250	NADH2 dehydrogenas
979	24	60.0	162	2	T14248	NADH2 dehydrogenas
980	24	60.0	162	2	T14251	NADH2 dehydrogenas
981	24	60.0	162	2	T14249	NADH2 dehydrogenas
982	24	60.0	162	2	T17042	NADH2 dehydrogenas
983	24	60.0	162	2	T14257	NADH2 dehydrogenas

984	24	60.0	162	2	T14252	NADH2 dehydrogenas
985	24	60.0	162	2	T17034	NADH2 dehydrogenas
986	24	60.0	162	2	T17038	NADH2 dehydrogenas
987	24	60.0	162	2	T17043	NADH2 dehydrogenas
988	24	60.0	162	2	D97668	probable transcrip
989	24	60.0	164	2	T07759	disease resistance
990	24	60.0	164	2	A81332	hypothetical prote
991	24	60.0	166	2	E71040	hypothetical prote
992	24	60.0	167	2	T47385	hypothetical prote
993	24	60.0	167	2	S53747	Ca(2+)-ATPase - ra
994	24	60.0	168	2	H90399	hypothetical prote
995	24	60.0	174	2	T38498	hypothetical prote
996	24	60.0	175	2	JN0682	gamma3-crystallin
997	24	60.0	175	2	JN0681	gamma2-crystallin
998	24	60.0	176	1	I64161	cytochrome c bioge
999	24	60.0	177	2	C83495	probable lipoprote
1000	24	60.0	177	2	JE0226	allergen Mal f2 -

# ALIGNMENTS

## RESULT 1

S23094

beta-amyloid protein precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 03-May-1996

C;Accession: S23094

R;Kojima, S.; Omori, M.

FEBS Lett. 304, 57-60, 1992

A;Title: Two-way cleavage of beta-amyloid protein precursor by multicatalytic proteinase.

A;Reference number: S23094; MUID:92316198; PMID:1618299

A;Accession: S23094

A;Molecule type: protein

A;Residues: 1-33 <KOJ>

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

Query Match 100.0%; Score 40; DB 2; Length 33;

Best Local Similarity 100.0%; Pred. No. 0.045;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8

|||||||

Db 21 KLVFFAED 28

## RESULT 2

PN0512

beta-amyloid protein - guinea pig (fragment)

C;Species: Cavia porcellus (guinea pig)

C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 17-Mar-1999

C;Accession: PN0512

R;Shimohigashi, Y.; Matsumoto, H.; Takano, Y.; Saito, R.; Iwata, T.; Kamiya, H.;

Ohno, M.

Biochem. Biophys. Res. Commun. 193, 624-630, 1993

A;Title: Receptor-mediated specific biological activity of a beta-amyloid protein fragment for NK-1 substance p receptors.  
A;Reference number: PN0512; MUID:93290653; PMID:7685598  
A;Accession: PN0512  
A;Molecule type: protein  
A;Residues: 1-42 <SHI>  
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology  
C;Keywords: alternative splicing; amyloid

Query Match 100.0%; Score 40; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.057;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 16 KLVFFAED 23

RESULT 3

E60045

Alzheimer's disease amyloid beta/A4 protein precursor - sheep (fragment)

C;Species: Ovis sp. (sheep)

C;Date: 01-Dec-1992 #sequence\_revision 01-Dec-1992 #text\_change 28-Jul-1995

C;Accession: E60045

R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis.

A;Reference number: A60045; MUID:92017079; PMID:1656157

A;Accession: E60045

A;Molecule type: mRNA

A;Residues: 1-57 <JOH>

A;Cross-references: EMBL:X56130

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 40; DB 2; Length 57;  
Best Local Similarity 100.0%; Pred. No. 0.078;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 21 KLVFFAED 28

RESULT 4

F60045

Alzheimer's disease amyloid beta/A4 protein precursor - pig (fragment)

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 01-Dec-1992 #sequence\_revision 01-Dec-1992 #text\_change 13-Aug-1999

C;Accession: F60045

R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis.

A;Reference number: A60045; MUID:92017079; PMID:1656157

A;Accession: F60045

A;Molecule type: mRNA

A;Residues: 1-57 <JOH>

A;Cross-references: EMBL:X56127; NID:g1895; PIDN:CAA39592.1; PID:g1896

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 40; DB 2; Length 57;  
Best Local Similarity 100.0%; Pred. No. 0.078;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 21 KLVFFAED 28

#### RESULT 5

G60045

Alzheimer's disease amyloid beta/A4 protein precursor - guinea pig (fragment)

C;Species: Cavia porcellus (guinea pig)

C;Date: 01-Dec-1992 #sequence\_revision 01-Dec-1992 #text\_change 28-Jul-1995

C;Accession: G60045

R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis.

A;Reference number: A60045; MUID:92017079; PMID:1656157

A;Accession: G60045

A;Molecule type: mRNA

A;Residues: 1-57 <JOH>

A;Cross-references: EMBL:X56126

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 40; DB 2; Length 57;  
Best Local Similarity 100.0%; Pred. No. 0.078;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 21 KLVFFAED 28

#### RESULT 6

D60045

Alzheimer's disease amyloid beta/A4 protein precursor - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 01-Dec-1992 #sequence\_revision 01-Dec-1992 #text\_change 28-Jul-1995

C;Accession: D60045

R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis.

A;Reference number: A60045; MUID:92017079; PMID:1656157

A;Accession: D60045

A;Molecule type: mRNA

A;Residues: 1-57 <JOH>

A;Cross-references: EMBL:X56124

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 40; DB 2; Length 57;  
Best Local Similarity 100.0%; Pred. No. 0.078;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 21 KLVFFAED 28

#### RESULT 7

A60045

Alzheimer's disease amyloid beta/A4 protein precursor - dog (fragment)

C;Species: Canis lupus familiaris (dog)

C;Date: 01-Dec-1992 #sequence\_revision 01-Dec-1992 #text\_change 28-Jul-1995

C;Accession: A60045

R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis.

A;Reference number: A60045; MUID:92017079; PMID:1656157

A;Accession: A60045

A;Molecule type: mRNA

A;Residues: 1-57 <JOH>

A;Cross-references: EMBL:X56125

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 40; DB 2; Length 57;  
Best Local Similarity 100.0%; Pred. No. 0.078;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 21 KLVFFAED 28

#### RESULT 8

B60045

Alzheimer's disease amyloid beta/A4 protein precursor - polar bear (fragment)

C;Species: Ursus maritimus (polar bear)

C;Date: 01-Dec-1992 #sequence\_revision 01-Dec-1992 #text\_change 13-Aug-1999

C;Accession: B60045



R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.  
 Brain Res. Mol. Brain Res. 10, 299-305, 1991  
 A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide  
 in dog, polar bear and five other mammals by cross-species polymerase chain  
 reaction analysis.  
 A;Reference number: A60045; MUID:92017079; PMID:1656157  
 A;Accession: B60045  
 A;Molecule type: mRNA  
 A;Residues: 1-57 <JOH>  
 A;Cross-references: EMBL:X56128; NID:g2165; PIDN:CAA39593.1; PID:g2166  
 C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type  
 proteinase inhibitor homology  
 C;Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 40; DB 2; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 0.078;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 21 KLVFFAED 28

#### RESULT 9

PQ0438

Alzheimer's disease amyloid A4 protein precursor - rabbit (fragment)

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 30-Sep-1993 #sequence\_revision 19-Oct-1995 #text\_change 19-Oct-1995

C;Accession: PQ0438; C60045

R;Davidson, J.S.; West, R.L.; Kotikalapudi, P.; Maroun, L.E.

Biochem. Biophys. Res. Commun. 188, 905-911, 1992

A;Title: Sequence and methylation in the beta/A4 region of the rabbit amyloid  
 precursor protein gene.

A;Reference number: PQ0438; MUID:93075180; PMID:1445331

A;Accession: PQ0438

A;Molecule type: DNA

A;Residues: 1-82 <DAV>

A;Cross-references: GB:M83558; GB:M83657

R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide  
 in dog, polar bear and five other mammals by cross-species polymerase chain  
 reaction analysis.

A;Reference number: A60045; MUID:92017079; PMID:1656157

A;Accession: C60045

A;Molecule type: mRNA

A;Residues: 12-68 <JOH>

A;Cross-references: EMBL:X56129

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type  
 proteinase inhibitor homology

C;Keywords: alternative splicing; Alzheimer's disease; amyloid; Down's syndrome

Query Match 100.0%; Score 40; DB 2; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8

|||||||  
Db 32 KLVFFAED 39

RESULT 10

A49795

Alzheimer's disease amyloid beta protein precursor - crab-eating macaque

C;Species: *Macaca fascicularis* (crab-eating macaque)

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C;Accession: A49795

R;Podlisny, M.B.; Tolan, D.R.; Selkoe, D.J.

Am. J. Pathol. 138, 1423-1435, 1991

A;Title: Homology of the amyloid beta protein precursor in monkey and human supports a primate model for beta amyloidosis in Alzheimer's disease.

A;Reference number: A49795; MUID:91273117; PMID:1905108

A;Accession: A49795

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-695 <POD>

A;Cross-references: GB:M58727; NID:g342062; PIDN:AAA36829.1; PID:g342063

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing

Query Match 100.0%; Score 40; DB 1; Length 695;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 612 KLVFFAED 619

RESULT 11

A27485

Alzheimer's disease amyloid beta/A4 protein homolog precursor - mouse

N;Alternate names: proteinase nexin II

C;Species: *Mus musculus* (house mouse)

C;Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 13-Aug-1999

C;Accession: A27485; S19727; I49485

R;Yamada, T.; Sasaki, H.; Furuya, H.; Miyata, T.; Goto, I.; Sakaki, Y.

Biochem. Biophys. Res. Commun. 149, 665-671, 1987

A;Title: Complementary DNA for the mouse homolog of the human amyloid beta protein precursor.

A;Reference number: A27485; MUID:88106489; PMID:3322280

A;Accession: A27485

A;Molecule type: mRNA

A;Residues: 1-695 <YAM>

A;Cross-references: GB:M18373; NID:g191568; PIDN:AAA37139.1; PID:g309085

A;Experimental source: brain

R;de Strooper, B.; van Leuven, F.; van den Berghe, H.

Biochim. Biophys. Acta 1129, 141-143, 1991

A;Title: The amyloid beta protein precursor or proteinase nexin II from mouse is closer related to its human homolog than previously reported.

A;Reference number: S19727; MUID:92096458; PMID:1756177

A;Accession: S19727

A;Molecule type: mRNA

A;Residues: 1-210,'G',212-220,'S',222-396,'A',398-402,'T',404-448,'A',450-695  
<STR>

A;Cross-references: EMBL:X59379

R;Izumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, H.; Hattori, M.; Sakaki, Y.  
Gene 112, 189-195, 1992

A;Title: Positive and negative regulatory elements for the expression of the  
Alzheimer's disease amyloid precursor-encoding gene in mouse.

A;Reference number: I49485; MUID:92209998; PMID:1555768

A;Accession: I49485

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-19 <RES>

A;Cross-references: GB:D10603; NID:g220328; PIDN:BAA01456.1; PID:g220329

C;Genetics:

A;Map position: 16C3

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type  
proteinase inhibitor homology

C;Keywords: alternative splicing; amyloid; transmembrane protein

Query Match 100.0%; Score 40; DB 2; Length 695;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
| | | | | | | |  
Db 612 KLVFFAED 619

#### RESULT 12

S00550

Alzheimer's disease amyloid beta protein precursor - rat

N;Alternate names: beta-A4 amyloid protein

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 13-Aug-1999

C;Accession: S00550; A41245; A39820; S46251

R;Shivers, B.D.; Hilbich, C.; Multhaup, G.; Salbaum, M.; Beyreuther, K.;  
Seeburg, P.H.

EMBO J. 7, 1365-1370, 1988

A;Title: Alzheimer's disease amyloidogenic glycoprotein: expression pattern in  
rat brain suggests a role in cell contact.

A;Reference number: S00550; MUID:88312583; PMID:2900758

A;Accession: S00550

A;Molecule type: mRNA

A;Residues: 1-695 <SHI>

A;Cross-references: EMBL:X07648; NID:g55616; PIDN:CAA30488.1; PID:g55617

R;Schubert, D.; Schroeder, R.; LaCorbiere, M.; Saitoh, T.; Cole, G.  
Science 241, 223-226, 1988

A;Title: Amyloid beta protein precursor is possibly a heparan sulfate  
proteoglycan core protein.

A;Reference number: A41245; MUID:88264430; PMID:2968652

A;Accession: A41245

A;Molecule type: protein

A;Residues: 18-37,'X',39-40,'X',42-44 <SCH>

A;Note: evidence for heparan sulfate attachment

R;Hesse, L.; Beher, D.; Masters, C.L.; Multhaup, G.

FEBS Lett. 349, 109-116, 1994

A;Title: The beta-A4 amyloid precursor protein binding to copper.

A;Reference number: S46251; MUID:94320627; PMID:7913895  
 A;Contents: annotation; copper binding sites  
 A;Note: rat peptides were isolated but not sequenced  
 R;Potempska, A.; Styles, J.; Mehta, P.; Kim, K.S.; Miller, D.L.  
 J. Biol. Chem. 266, 8464-8469, 1991  
 A;Title: Purification and tissue level of the beta-amyloid peptide precursor of rat brain.  
 A;Reference number: A39820; MUID:91217087; PMID:1673681  
 A;Accession: A39820  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 18-32 <POT>  
 A;Experimental source: brain  
 C;Comment: Deposition of amyloid protein as neurofibrillary tangles and/or plaques is characteristic of both Alzheimer's disease and Down's syndrome.  
 C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology  
 C;Keywords: alternative splicing; amyloid; glycoprotein; transmembrane protein  
 F;625-648/Domain: transmembrane #status predicted <TMM>

Query Match 100.0%; Score 40; DB 2; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 612 KLVFFAED 619

# RESULT 13

## QRHUA4

Alzheimer's disease amyloid beta protein precursor [validated] - human  
 N;Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor XIa inhibitor; proteinase nexin II (PN-II)  
 N;Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascular form; amyloid protein precursor splice form APP(695); amyloid protein precursor splice form APP(751); amyloid protein precursor splice form APP(770)  
 C;Species: Homo sapiens (man)  
 C;Date: 30-Jun-1987 #sequence\_revision 28-Jul-1995 #text\_change 15-Sep-2000  
 C;Accession: S02260; S05194; A32277; A33260; A35486; I39452; I39451; I39453; I59562; A44017; B44017; A03134; A29030; A47584; A47585; S02638; S00707; S00925; A38949; A30320; B30320; C30320; A31087; A24668; A28583; A29302; A60805; JL0038; S06121; A60355; A59011; A38384; S29076; S38252; S32539; S48148; S48692; S51186; S51185; S51184; S51183; A54238; I58075; I52250; S09010; S10737; S24127; S43644  
 R;Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayney, R.M.; Unterbeck, A.; Beyreuther, K.; Mueller-Hill, B.  
 Nucleic Acids Res. 17, 517-522, 1989  
 A;Title: The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid is encoded by 16 exons.  
 A;Reference number: S02260; MUID:89128427; PMID:2783775  
 A;Accession: S02260  
 A;Molecule type: DNA  
 A;Residues: 1-288, 'V', 365-770 <LEM1>  
 A;Cross-references: EMBL:X13466  
 A;Note: alternative splice form APP(695)  
 R;Lemaire, H.G.  
 submitted to the EMBL Data Library, November 1988

A;Reference number: S05194  
 A;Accession: S05194  
 A;Molecule type: DNA  
 A;Residues: 1-14,'VW',17-288,'V',365-770 <LEM2>  
 A;Cross-references: EMBL:X13466; NID:g35598; PIDN:CAA31830.1; PID:g871360  
 A;Note: alternative splice form APP(695)  
 R;La Fauci, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.  
 Biochem. Biophys. Res. Commun. 159, 297-304, 1989  
 A;Title: Characterization of the 5'-end region and the first two exons of the beta-protein precursor gene.  
 A;Reference number: A32277; MUID:89165870; PMID:2538123  
 A;Accession: A32277  
 A;Molecule type: DNA  
 A;Residues: 1-75 <LAF>  
 A;Cross-references: GB:M24546; GB:M24547; NID:g341202; PIDN:AAC13654.1; PID:g516074  
 R;Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little, S.P.  
 Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989  
 A;Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows similarity to soybean trypsin inhibitor.  
 A;Reference number: A33260; MUID:89392030; PMID:2675837  
 A;Accession: A33260  
 A;Molecule type: DNA  
 A;Residues: 656-737 <JOH>  
 A;Cross-references: GB:M29270; NID:g178863; PIDN:AAA51768.1; PID:g178865  
 R;Prelli, F.; Levy, E.; van Duinen, S.G.; Bots, G.T.A.M.; Luyendijk, W.; Frangione, B.  
 Biochem. Biophys. Res. Commun. 170, 301-307, 1990  
 A;Title: Expression of a normal and variant Alzheimer's beta-protein gene in amyloid of hereditary cerebral hemorrhage, Dutch type: DNA and protein diagnostic assays.  
 A;Reference number: A35486; MUID:90321244; PMID:2196878  
 A;Accession: A35486  
 A;Molecule type: DNA  
 A;Residues: 672-710 <PRE1>  
 A;Note: 693-Gln was found in DNA isolated from HCHWA-D patients  
 R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.  
 Gene 87, 257-263, 1990  
 A;Title: Genomic organization of the human amyloid beta-protein precursor gene.  
 A;Reference number: I39451; MUID:90236318; PMID:2110105  
 A;Accession: I39452  
 A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-770 <YOS1>  
 A;Cross-references: GB:M33112; NID:g178613; PIDN:AAB59502.1; PID:g178616  
 A;Accession: I39451  
 A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-530,'QWLMPVIPAFWEAKVGR' <YOS2>  
 A;Cross-references: GB:M34875; NID:g178608; PIDN:AAB59501.1; PID:g178615  
 R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.  
 Gene 102, 291-292, 1991  
 A;Reference number: A59020; MUID:91340168; PMID:1908403  
 A;Contents: annotation; erratum

A;Note: revised physical map for reference I39451  
 R;Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.;  
 van Duinen, S.G.; Bots, G.T.; Luyendijk, W.; Frangione, B.  
 Science 248, 1124-1126, 1990  
 A;Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral  
 hemorrhage, Dutch type.  
 A;Reference number: I39453; MUID:90260663; PMID:2111584  
 A;Accession: I39453  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 656-737 <LEV>  
 A;Cross-references: GB:M37896; NID:g178618; PIDN:AAA51727.1; PID:g178620  
 A;Note: a mutation with 693-Gln is presented  
 R;Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.  
 Science 254, 97-99, 1991  
 A;Title: A mutation in the amyloid precursor protein associated with hereditary  
 Alzheimer's disease.  
 A;Reference number: I59562; MUID:92022553; PMID:1925564  
 A;Accession: I59562  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 689-716,'F',718-737 <MUR>  
 A;Cross-references: GB:S57665; NID:g236720; PIDN:AAB19991.1; PID:g236721  
 R;Kamino, K.; Orr, H.T.; Payami, H.; Wijsman, E.M.; Alonso, M.E.; Pulst, S.M.;  
 Anderson, L.; O'dahl, S.; Nemens, E.; White, J.A.; Sadovnick, A.D.; Ball, M.J.;  
 Kaye, J.; Warren, A.; McInnis, M.; Antonarakis, S.E.; Korenberg, J.R.; Sharma,  
 V.; Kukull, W.; Larson, E.; Heston, L.L.; Martin, G.M.; Bird, T.D.;  
 Schellenberg, G.D.  
 Am. J. Hum. Genet. 51, 998-1014, 1992  
 A;Title: Linkage and mutational analysis of familial Alzheimer disease kindreds  
 for the APP gene region.  
 A;Reference number: A44017; MUID:93035397; PMID:1415269  
 A;Accession: A44017  
 A;Molecule type: DNA  
 A;Residues: 687-692,'G',694-718 <KAM1>  
 A;Cross-references: GB:S45135; NID:g257377; PIDN:AAB23645.1; PID:g257378  
 A;Experimental source: familial Alzheimer disease family SB  
 A;Note: sequence extracted from NCBI backbone (NCBIP:115374)  
 A;Accession: B44017  
 A;Molecule type: DNA  
 A;Residues: 687-718 <KAM2>  
 A;Cross-references: GB:S45136; NID:g257379; PIDN:AAB23646.1; PID:g257380  
 A;Experimental source: familial Alzheimer disease family LIT  
 A;Note: sequence extracted from NCBI backbone (NCBIP:115376)  
 A;Note: this sequence has a silent mutation  
 R;Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.;  
 Grzeschik, K.H.; Multhaup, G.; Beyreuther, K.; Muller-Hill, B.  
 Nature 325, 733-736, 1987  
 A;Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a  
 cell-surface receptor.  
 A;Reference number: A03134; MUID:87144572; PMID:2881207  
 A;Accession: A03134  
 A;Molecule type: mRNA  
 A;Residues: 1-288,'V',365-770 <KAN>  
 A;Cross-references: GB:Y00264; NID:g28525; PIDN:CAA68374.1; PID:g28526  
 A;Note: alternative splice form APP(695)  
 R;Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.

Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987  
 A;Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascular and the neuritic plaque amyloid peptides.  
 A;Reference number: A29030; MUID:87231971; PMID:3035574  
 A;Accession: A29030  
 A;Molecule type: mRNA  
 A;Residues: 284-288,'V',365-646,'E',648-770 <ROB>  
 A;Cross-references: GB:M16765; NID:g178539; PIDN:AAA51722.1; PID:g178540  
 A;Note: the authors translated the codon GAG for residue 647 as Asp  
 R;Goldgaber, D.; Lerman, M.I.; McBride, O.W.; Saffiotti, U.; Gajdusek, D.C.  
 Science 235, 877-880, 1987  
 A;Title: Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease.  
 A;Reference number: A47584; MUID:87120328; PMID:3810169  
 A;Accession: A47584  
 A;Molecule type: mRNA  
 A;Residues: 674-756,'S',758-770 <GOL>  
 A;Cross-references: GB:M15533; NID:g178706; PIDN:AAA35540.1; PID:g178707  
 A;Experimental source: brain  
 R;Tanzi, R.E.; Gusella, J.F.; Watkins, P.C.; Bruns, G.A.P.; St George-Hyslop, P.; Van Keuren, M.L.; Patterson, D.; Pagan, S.; Kurnit, D.M.; Neve, R.L.  
 Science 235, 880-884, 1987  
 A;Title: Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus.  
 A;Reference number: A47585; MUID:87120329; PMID:2949367  
 A;Accession: A47585  
 A;Molecule type: mRNA  
 A;Residues: 674-703 <TAN1>  
 A;Cross-references: GB:M15532; NID:g177957; PIDN:AAA51564.1; PID:g177958  
 R;Dyrks, T.; Weidemann, A.; Multhaup, G.; Salbaum, J.M.; Lemaire, H.G.; Kang, J.; Mueller-Hill, B.; Masters, C.L.; Beyreuther, K.  
 EMBO J. 7, 949-957, 1988  
 A;Title: Identification, transmembrane orientation and biogenesis of the amyloid A4 precursor of Alzheimer's disease.  
 A;Reference number: S02638; MUID:88296437; PMID:2900137  
 A;Accession: S02638  
 A;Molecule type: mRNA  
 A;Residues: 672-678 <DYR>  
 R;Tanzi, R.E.; McClatchey, A.I.; Lamperti, E.D.; Villa-Komaroff, L.; Gusella, J.F.; Neve, R.L.  
 Nature 331, 528-530, 1988  
 A;Title: Protease inhibitor domain encoded by an amyloid protein precursor mRNA associated with Alzheimer's disease.  
 A;Reference number: S00707; MUID:88122640; PMID:2893290  
 A;Accession: S00707  
 A;Molecule type: mRNA  
 A;Residues: 286-344,'I',365-366 <TAN2>  
 A;Cross-references: EMBL:X06982; NID:g28817; PIDN:CAA30042.1; PID:g929612  
 A;Experimental source: promyelocytic leukemia cell line HL60  
 A;Note: alternative splice form APP(751)  
 R;Ponte, P.; Gonzalez-DeWhitt, P.; Schilling, J.; Miller, J.; Hsu, D.; Greenberg, B.; Davis, K.; Wallace, W.; Lieberburg, I.; Fuller, F.; Cordell, B.  
 Nature 331, 525-527, 1988  
 A;Title: A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibitors.  
 A;Reference number: S00925; MUID:88122639; PMID:2893289  
 A;Accession: S00925

A;Molecule type: mRNA  
 A;Residues: 1-344,'I',365-770 <PO2>  
 A;Cross-references: GB:X06989; EMBL:Y00297; NID:g28720; PIDN:CAA30050.1; PID:g28721  
 A;Note: alternative splice form APP(751)  
 R;Kitaguchi, N.; Takahashi, Y.; Tokushima, Y.; Shiojiri, S.; Ito, H.  
 Nature 331, 530-532, 1988  
 A;Title: Novel precursor of Alzheimer's disease amyloid protein shows protease inhibitory activity.  
 A;Reference number: A38949; MUID:88122641; PMID:2893291  
 A;Accession: A38949  
 A;Molecule type: mRNA  
 A;Residues: 287-367 <KIT>  
 A;Cross-references: GB:X06981; NID:g28816; PIDN:CAA30041.1; PID:g929611  
 A;Experimental source: glioblastoma cell line  
 A;Note: alternative splice form APP(770)  
 R;Vitek, M.P.; Rasool, C.G.; de Sauvage, F.; Vitek, S.M.; Bartus, R.T.; Beer, B.; Ashton, R.A.; Macq, A.F.; Maloteaux, J.M.; Blume, A.J.; Octave, J.N.  
 Brain Res. Mol. Brain Res. 4, 121-131, 1988  
 A;Title: Absence of mutation in the beta-amyloid cDNAs cloned from the brains of three patients with sporadic Alzheimer's disease.  
 A;Reference number: A30320  
 A;Accession: A30320  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 284-288,'V',365-770 <VIT1>  
 A;Accession: B30320  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 122-288,'V',365-770 <VIT2>  
 A;Accession: C30320  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 606-770 <VIT3>  
 R;Zain, S.B.; Salim, M.; Chou, W.G.; Sajdel-Sulkowska, E.M.; Majocha, R.E.; Marotta, C.A.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 929-933, 1988  
 A;Title: Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer disease brain: coding and noncoding regions of the fetal precursor mRNA are expressed in the cortex.  
 A;Reference number: A31087; MUID:88124954; PMID:2893379  
 A;Accession: A31087  
 A;Molecule type: mRNA  
 A;Residues: 507-770 <ZAI>  
 A;Cross-references: GB:M18734; NID:g178572; PIDN:AAA51726.1; PID:g178573  
 A;Note: the authors translated the codon GAA for residue 599 as Gly, ACC for residue 603 as Val, GTG for residue 604 as Glu, GAG for residue 605 as Leu, CTT for residue 607 as Pro, CCC for residue 608 as Val, GTG for residue 609 as Asn, AAT for residue 610 as Gly, and GGT for residue 655 as Ser  
 A;Note: the cited Genbank accession number, J03594, is not in release 101.0  
 R;Masters, C.L.; Multhaup, G.; Simms, G.; Pottgiesser, J.; Martins, R.N.; Beyreuther, K.

Query Match 100.0%; Score 40; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy           1 KLVFFAED 8  
              |||||||  
Db           687 KLVFFAED 694

RESULT 14

JH0773

Alzheimer's disease amyloid beta protein precursor - African clawed frog

C;Species: *Xenopus laevis* (African clawed frog)

C;Date: 10-Jun-1993 #sequence\_revision 10-Jun-1993 #text\_change 13-Aug-1999

C;Accession: JH0773

R;Okado, H.; Okamoto, H.

Biochem. Biophys. Res. Commun. 189, 1561-1568, 1992

A;Title: A *Xenopus* homologue of the human beta-amyloid precursor protein: developmental regulation of its gene expression.

A;Reference number: JH0773; MUID:93129227; PMID:1282805

A;Accession: JH0773

A;Molecule type: mRNA

A;Residues: 1-747 <OKA>

A;Cross-references: GB:S52417; NID:g263150; PIDN:AAB24853.1; PID:g263151

A;Experimental source: larva

C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology

C;Keywords: alternative splicing; amyloid

F;287-337/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

Query Match                   90.0%; Score 36; DB 2; Length 747;  
Best Local Similarity       87.5%; Pred. No. 8.1;  
Matches       7; Conservative   1; Mismatches       0; Indels       0; Gaps       0;

Qy           1 KLVFFAED 8  
              |||||||:  
Db           664 KLVFFAEE 671

RESULT 15

H71729

hypothetical protein RP189 - *Rickettsia prowazekii*

C;Species: *Rickettsia prowazekii*

C;Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000

C;Accession: H71729

R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.;

Alsmark, U.C.M.; Podowski, R.M.; Naeslund, A.K.; Eriksson, A.S.; Winkler, H.H.;

Kurland, C.G.

Nature 396, 133-140, 1998

A;Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.

A;Reference number: A71630; MUID:99039499; PMID:9823893

A;Accession: H71729

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-321 <AND>

A;Cross-references: GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAA14655.1;

PID:e1342498; PID:g3860754; GSPDB:GN00081

A;Experimental source: strain Madrid E

C;Genetics:

A;Gene: RP189

C;Superfamily: Rickettsia prowazekii hypothetical protein RP189

Query Match 85.0%; Score 34; DB 2; Length 321;  
Best Local Similarity 75.0%; Pred. No. 9.5;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||:|||||  
Db 178 KLIFFAHD 185

#### RESULT 16

T35807

hypothetical protein SC8D9.03 SC8D9.03 - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jan-2000

C;Accession: T35807

R;Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.;  
Rajandream, M.A.

submitted to the EMBL Data Library, February 1999

A;Reference number: Z21589

A;Accession: T35807

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-182 <MUR>

A;Cross-references: EMBL:AL035569; PIDN:CAB37567.1; GSPDB:GN00070;

SCOEDB:SC8D9.03

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SC8D9.03

C;Superfamily: yeast conserved hypothetical protein YJL055w

Query Match 80.0%; Score 32; DB 2; Length 182;  
Best Local Similarity 75.0%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
:|||||:  
Db 160 ELVFFAEE 167

#### RESULT 17

B89868

conserved hypothetical protein SA0860 [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C;Accession: B89868

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Aoki, K.; Nagai, Y.; Lian, J.; Ito, T.; Kanamori, M.; Matsumaru, H.; Maruyama, A.; Murakami, H.; Hosoyama, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Hirakawa, H.; Kuhara, S.; Goto, S.; Yabuzaki, J.; Kanehisa, M.; Yamashita, A.; Oshima, K.; Furuya, K.; Yoshino, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: B89868  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-261 <KUR>  
A;Cross-references: GB:BA000018; PID:g13700805; PIDN:BAB42101.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: SA0860

Query Match 80.0%; Score 32; DB 2; Length 261;  
Best Local Similarity 85.7%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
||||:|  
Db 158 LVFFSED 164

RESULT 18

D69345

LSU ribosomal protein L7AE (rpl7AE) homolog - *Archaeoglobus fulgidus*

C;Species: *Archaeoglobus fulgidus*

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 13-Aug-1999

C;Accession: D69345

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.; Peterson, J.D.; Richardson, D.L.; Kerlavage, A.R.; Graham, D.E.; Kyrpides, N.C.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.; Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.; D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.; Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.  
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon *Archaeoglobus fulgidus*.

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: D69345

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-119 <KLE>

A;Cross-references: GB:AE001051; GB:AE000782; NID:g2689374; PIDN:AAB90466.1;

PID:g2649836; TIGR:AF0764

C;Superfamily: rat ribosomal protein L7a

Query Match 77.5%; Score 31; DB 2; Length 119;  
Best Local Similarity 75.0%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||: |||  
Db 45 KLVYIAED 52

RESULT 19

H69651

lichenan operon transcription antiterminator licR - *Bacillus subtilis*  
 N;Alternate names: cel operon regulator  
 C;Species: *Bacillus subtilis*  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
 C;Accession: H69651; S57758  
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Guiseppe, G.; Guy, B.J.; Haga, K.; Haiech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.  
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwollik, S.; Prescott, A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.  
 A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.; Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenegger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.  
 A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377  
 A;Accession: H69651  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-641 <KUN>  
 A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15886.1; PID:ell86359; PID:g2636395  
 A;Experimental source: strain 168  
 R;Glaser, P.; Lubochinsky, B.; Danchin, A.  
 submitted to the EMBL Data Library, July 1995  
 A;Description: *Bacillus subtilis* cel operon.  
 A;Reference number: S57758  
 A;Accession: S57758  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-641 <GLA>  
 A;Cross-references: EMBL:Z49992; NID:g895746; PIDN:CAA90284.1; PID:g895747  
 C;Genetics:  
 A;Gene: licR  
 C;Keywords: transcription antitermination

Query Match 77.5%; Score 31; DB 2; Length 641;  
Best Local Similarity 75.0%; Pred. No. 88;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
:|||| ||  
Db 504 ELVFFQED 511

RESULT 20

T38991

conserved hypothetical protein SPAC630.13c - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C;Accession: T38991

R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Zimmermann, W.; Wambutt, R.  
submitted to the EMBL Data Library, August 1999

A;Reference number: Z21822

A;Accession: T38991

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1339 <MCD>

A;Cross-references: EMBL:AL109832; PIDN:CAB52735.1; GSPDB:GN00066;

SPDB:SPAC630.13c

A;Experimental source: strain 972h-; cosmid c630

C;Genetics:

A;Gene: SPDB:SPAC630.13c

A;Map position: 1

Query Match 77.5%; Score 31; DB 2; Length 1339;  
Best Local Similarity 75.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||||| :|  
Db 419 KLVFFLKD 426

RESULT 21

T51920

probable xanthine dehydrogenase [imported] - *Neurospora crassa*

N;Alternate names: protein B23I11.320

C;Species: *Neurospora crassa*

C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 01-Dec-2000

C;Accession: T51920

R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.;  
Nyakatura, G.; Mewes, H.W.; Mannhaupt, G.

submitted to the Protein Sequence Database, August 2000

A;Reference number: Z25858

A;Accession: T51920

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1364 <SCH>

A;Cross-references: EMBL:AL391572; GSPDB:GN00116; NCSP:B23I11.320

A;Experimental source: BAC clone B23I11; strain OR74A

C;Genetics:  
A;Gene: NCSP:B23I11.320  
A;Map position: 6  
A;Introns: 66/2; 1321/3  
C;Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology  
C;Keywords: 2Fe-2S; metalloprotein  
F;68,73,76,98/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 77.5%; Score 31; DB 2; Length 1364;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
| | | | |  
Db 681 VFFAED 686

#### RESULT 22

T24151  
hypothetical protein R10H10.1 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T24151  
R;Bardill, S.  
submitted to the EMBL Data Library, April 1996  
A;Reference number: Z19846  
A;Accession: T24151  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-222 <WIL>  
A;Cross-references: EMBL:Z70686; PIDN:CAA94609.1; GSPDB:GN00022; CESP:R10H10.1  
A;Experimental source: clone R10H10  
C;Genetics:  
A;Gene: CESP:R10H10.1  
A;Map position: 4  
A;Introns: 13/1; 34/1; 60/2

Query Match 75.0%; Score 30; DB 2; Length 222;  
Best Local Similarity 75.0%; Pred. No. 49;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | |  
Db 69 KRVFFGED 76

#### RESULT 23

AG0459  
Sec-independent protein translocase protein TatC [imported] - *Yersinia pestis* (strain CO92)  
C;Species: *Yersinia pestis*  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
C;Accession: AG0459  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; Sebahia, M.; James, K.D.; Churcher, C.; Mungall, K.L.; Baker, S.; Basham, D.; Bentley, S.D.; Brooks, K.; Cerdeno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Feltwell, T.; Hamlin, N.;

Holroyd, S.; Jagels, K.; Leather, S.; Karlyshev, A.V.; Moule, S.; Oyston, P.C.F.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.

Nature 413, 523-527, 2001

A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AG0459

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-258 <KUR>

A;Cross-references: GB:AL590842; PIDN:CAC93243.1; PID:g15981689; GSPDB:GN00175

C;Genetics:

A;Gene: tatC

C;Superfamily: conserved hypothetical protein HI0188

Query Match 75.0%; Score 30; DB 2; Length 258;

Best Local Similarity 85.7%; Pred. No. 58;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8

|||||

Db 34 LVFFAND 40

#### RESULT 24

A64383

hypothetical protein MJ0665 - *Methanococcus jannaschii*

C;Species: *Methanococcus jannaschii*

C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C;Accession: A64383

R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton, R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.; Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann, J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*.

A;Reference number: A64300; MUID:96337999; PMID:8688087

A;Accession: A64383

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-341 <BUL>

A;Cross-references: GB:U67513; GB:L77117; NID:g1591365; PIDN:AAB98656.1;

PID:g1591378; TIGR:MJ0665

C;Genetics:

A;Map position: REV591204-590179

A;Start codon: GTG

C;Superfamily: hypothetical protein MJ0665

Query Match 75.0%; Score 30; DB 2; Length 341;

Best Local Similarity 71.4%; Pred. No. 76;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
||:|:|  
Db 67 KLIFYAE 73

RESULT 25

T47131

G-protein coupled receptor, SREB2 - human

C;Species: Homo sapiens (man)

C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 21-Jun-2002

C;Accession: T47131; JC7288

R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.  
submitted to the Protein Sequence Database, March 2000

A;Reference number: Z24374

A;Accession: T47131

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-370 <AAA>

A;Cross-references: EMBL:AL161959; NID:g7328012; PIDN:CAB82307.1; PID:g7328013

A;Experimental source: adult amygdala; clone DKFZp761L08121

R;Matsumoto, M.; Saito, T.; Takasaki, J.; Kamohara, M.; Sugimoto, T.; Kobayashi, M.; Tadokoro, M.; Matsumoto, S.; Ohishi, T.; Furuichi, K.

Biochem. Biophys. Res. Commun. 272, 576-582, 2000

A;Title: An evolutionarily conserved G-protein coupled receptor family, SREB, expressed in the central nervous system.

A;Reference number: JC7287

A;Accession: JC7288

A;Molecule type: mRNA

A;Residues: 1-370 <MAT>

A;Cross-references: DDBJ:AB040799

C;Genetics:

A;Gene: sreb2

A;Map position: 7q31

A;Note: DKFZp761L08121.1

C;Keywords: brain; G protein-coupled receptor; glycolysis; reproduction; transmembrane protein

Query Match 75.0%; Score 30; DB 2; Length 370;  
Best Local Similarity 62.5%; Pred. No. 83;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||:| |  
Db 204 KLIFVHD 211

RESULT 26

T27908

hypothetical protein ZK550.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T27908

R;Basham, V.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z20438

A;Accession: T27908



A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-502 <WIL>  
A;Cross-references: EMBL:Z82287; PIDN:CAB05312.1; GSPDB:GN00022; CESP:ZK550.2  
A;Experimental source: clone ZK550  
C;Genetics:  
A;Gene: CESP:ZK550.2  
A;Map position: 4  
A;Introns: 23/2; 88/3; 187/3; 247/1; 281/2; 328/2; 364/1; 446/2

Query Match 75.0%; Score 30; DB 2; Length 502;  
Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
|||| ||  
Db 227 LVFFQED 233

RESULT 27

T46975

lysine-tRNA ligase (EC 6.1.1.6) [validated] - Methanococcus maripaludis

N;Alternate names: lysyl-tRNA synthetase

C;Species: Methanococcus maripaludis

C;Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 03-Jun-2002

C;Accession: T46975

R;Ibba, M.; Morgan, S.; Curnow, A.W.; Pridmore, D.R.; Vothknecht, U.C.; Gardner, W.; Lin, W.; Woese, C.R.; Soll, D.

Science 278, 1119-1122, 1997

A;Title: A euryarchaeal lysyl-tRNA synthetase: resemblance to class I synthetases.

A;Reference number: A58679; MUID:98016282; PMID:9353192

A;Accession: T46975

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-533 <IBB>

A;Cross-references: EMBL:AF009824; NID:g2645488; PIDN:AAB87410.1; PID:g2645489

A;Experimental source: strain JJ

C;Genetics:

A;Gene: lysS

C;Function:

A;Description: EC 6.1.1.6 [validated, MUID:98016282]

C;Superfamily: Lyme disease spirochete lysine-tRNA ligase

C;Keywords: ligase

Query Match 75.0%; Score 30; DB 2; Length 533;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||:| |:|  
Db 58 KLIFIADD 65

RESULT 28

T39814

hypothetical protein SPBC19C7.08c - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C;Accession: T39814  
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gaillardin, C.  
submitted to the EMBL Data Library, June 1998  
A;Reference number: Z21881  
A;Accession: T39814  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-681 <LYN>  
A;Cross-references: EMBL:AL023859; PIDN:CAA19576.1; GSPDB:GN00067;  
SPDB:SPBC19C7.08c  
A;Experimental source: strain 972h-; cosmid c19C7  
C;Genetics:  
A;Gene: SPDB:SPBC19C7.08c  
A;Map position: 2

Query Match 75.0%; Score 30; DB 2; Length 681;  
Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|:| | | :|  
Db 589 KVVFFGDD 596

#### RESULT 29

S61568

probable membrane protein YDR205w - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein YD8142A.02

C;Species: Saccharomyces cerevisiae

C;Date: 09-Mar-1996 #sequence\_revision 12-Apr-1996 #text\_change 19-Apr-2002

C;Accession: S61568

R;Murphy, L.; Harris, D.

submitted to the EMBL Data Library, December 1995

A;Reference number: S61117

A;Accession: S61568

A;Molecule type: DNA

A;Residues: 1-740 <MUR>

A;Cross-references: EMBL:Z68194; NID:g1204148; PID:e213788; PID:g1122333;

GSPDB:GN00004; MIPS:YDR205w

A;Experimental source: strain AB972

C;Genetics:

A;Gene: SGD:MSC2; MIPS:YDR205w

A;Cross-references: SGD:S0002613

A;Map position: 4R

C;Keywords: transmembrane protein

F;22-38/Domain: transmembrane #status predicted <TM1>

F;107-123/Domain: transmembrane #status predicted <TM2>

F;155-171/Domain: transmembrane #status predicted <TM3>

F;196-212/Domain: transmembrane #status predicted <TM4>

F;239-255/Domain: transmembrane #status predicted <TM5>

F;264-280/Domain: transmembrane #status predicted <TM6>

F;318-334/Domain: transmembrane #status predicted <TM7>

F;346-362/Domain: transmembrane #status predicted <TM8>

F;404-420/Domain: transmembrane #status predicted <TM9>

F;437-453/Domain: transmembrane #status predicted <TM10>

F;477-493/Domain: transmembrane #status predicted <TM11>  
F;508-524/Domain: transmembrane #status predicted <TM12>  
F;544-560/Domain: transmembrane #status predicted <TM13>  
F;576-592/Domain: transmembrane #status predicted <TM14>

Query Match 75.0%; Score 30; DB 2; Length 740;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||:|:|  
Db 65 KLIFFSTD 72

RESULT 30

T45876

hypothetical protein F4P12.60 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 20-Jun-2000

C;Accession: T45876

R;Bloeker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23016

A;Accession: T45876

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-768 <BLO>

A;Cross-references: EMBL:AL132966

A;Experimental source: cultivar Columbia; BAC clone F4P12

C;Genetics:

A;Map position: 3

A;Note: F4P12.60

C;Superfamily: Arabidopsis thaliana hypothetical protein T8H10.30

Query Match 75.0%; Score 30; DB 2; Length 768;  
Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
:|||||  
Db 724 IFFAED 729

RESULT 31

JC4279

adenylate cyclase (EC 4.6.1.1) type 10 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 29-Nov-1995 #sequence\_revision 08-Feb-1996 #text\_change 26-Aug-1999

C;Accession: JC4279

R;Paterson, J.M.; Smith, S.M.; Harmar, A.J.; Antoni, F.A.

Biochem. Biophys. Res. Commun. 214, 1000-1008, 1995

A;Title: Control of a novel adenylyl cyclase by calcineurin.

A;Reference number: JC4279; MUID:96024597; PMID:7575502

A;Accession: JC4279

A;Molecule type: mRNA

A;Residues: 1-1353 <PAT>

A;Cross-references: EMBL:Z50190  
A;Experimental source: Att20 cells  
C;Comment: This enzyme is the target of regulation by calcineurin and is a physiologically relevant docking site for calcineurin. It converts ATP to cAMP, one of the earliest recognized intracellular messenger molecules.  
C;Superfamily: human adenylate cyclase; guanylate cyclase catalytic domain homology  
C;Keywords: phosphorus-oxygen lyase; transmembrane protein  
F;118-137/Domain: transmembrane #status predicted <TM01>  
F;142-162/Domain: transmembrane #status predicted <TM02>  
F;173-193/Domain: transmembrane #status predicted <TM03>  
F;230-241/Domain: transmembrane #status predicted <TM04>  
F;244-264/Domain: transmembrane #status predicted <TM05>  
F;284-304/Domain: transmembrane #status predicted <TM06>  
F;313-575/Domain: guanylate cyclase catalytic domain homology <GCC>  
F;787-807/Domain: transmembrane #status predicted <TM07>  
F;819-839/Domain: transmembrane #status predicted <TM08>  
F;865-885/Domain: transmembrane #status predicted <TM09>  
F;891-911/Domain: transmembrane #status predicted <TM10>  
F;921-941/Domain: transmembrane #status predicted <TM11>  
F;977-997/Domain: transmembrane #status predicted <TM12>  
F;1011-1246/Domain: guanylate cyclase catalytic domain homology <GCC2>

Query Match 75.0%; Score 30; DB 2; Length 1353;  
Best Local Similarity 62.5%; Pred. No. 3.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
::||| ||  
Db 840 RMVFFLED 847

# RESULT 32

F95064

ribosomal protein L7A family [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001

C;Accession: F95064

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidelberg, J.; DeBoy, R.T.; Haft, D.H.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.; Kolonay, J.F.; Nelson, W.C.; Peterson, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, E.; Khouri, H.; Wolf, A.M.; Utterback, T.R.; Hansen, C.L.; McDonald, L.A.; Feldblyum, T.V.; Angiuoli, S.; Dickinson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, D.A.; Hollingshead, S.K.; Fraser, C.M.

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: F95064

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-99 <KUR>

A;Cross-references: GB:AE005672; PIDN:AAK74711.1; PID:g14972029; GSPDB:GN00164; TIGR:SP4SP0555

A;Experimental source: strain TIGR4  
C;Genetics:  
A;Gene: SP0555

Query Match 72.5%; Score 29; DB 2; Length 99;  
Best Local Similarity 75.0%; Pred. No. 36;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||||||  
Db 36 KLVFLAHD 43

#### RESULT 33

H97931

conserved hypothetical protein spr0480 [imported] - *Streptococcus pneumoniae*  
(strain R6)

C;Species: *Streptococcus pneumoniae*

C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001

C;Accession: H97931

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Estrem, S.; Fritz, L.; Fu, D.J.; Fuller, W.; Geringer, C.; Gilmour, R.; Glass, J.S.; Khoja, H.; Kraft, A.; LaGace, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; McHenney, M.; McLeaster, K.; Mundy, C.; Nicas, T.I.; Norris, F.H.; O'Gara, M.; Peery, R.; Robertson, G.T.; Rockey, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.;

Jaskunas, S.R.; Rosteck Jr., P.R.; Skatrud, P.L.; Glass, J.I.

A;Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.

A;Reference number: A97872; MUID:21429245; PMID:11544234

A;Accession: H97931

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-99 <KUR>

A;Cross-references: GB:AE007317; PIDN:AAK99284.1; PID:g15458050; GSPDB:GN00174

C;Genetics:

A;Gene: spr0480

Query Match 72.5%; Score 29; DB 2; Length 99;  
Best Local Similarity 75.0%; Pred. No. 36;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||||||  
Db 36 KLVFLAHD 43

#### RESULT 34

AH1192

B. subtilis YneR protein homolog lmo0944 [imported] - *Listeria monocytogenes*  
(strain EGD-e)

C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Dec-2002

C;Accession: AH1192

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve,

E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.  
 A;Title: Comparative genomics of *Listeria* species.  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AH1192  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-100 <GLA>  
 A;Cross-references: GB:NC\_003210; PIDN:CAC99022.1; PID:g16410346; GSPDB:GN00177  
 A;Experimental source: strain EGD-e  
 C;Genetics:  
 A;Gene: lmo0944  
 C;Superfamily: uncharacterized conserved protein

Query Match 72.5%; Score 29; DB 2; Length 100;  
 Best Local Similarity 71.4%; Pred. No. 36;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
 |:|| ||  
 Db 63 LIFFIED 69

# RESULT 35

R5HSS6  
 ribosomal protein HS6 [validated] - *Haloarcula marismortui*  
 C;Species: *Haloarcula marismortui*  
 C;Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 21-Jul-2000  
 C;Accession: S00182  
 R;Kimura, J.; Arndt, E.; Kimura, M.  
 FEBS Lett. 224, 65-70, 1987  
 A;Title: Primary structures of three highly acidic ribosomal proteins S6, S12 and S15 from the archaebacterium *Halobacterium marismortui*.  
 A;Reference number: S00182; MUID:88055606; PMID:3315748  
 A;Accession: S00182  
 A;Molecule type: protein  
 A;Residues: 1-116 <KIM>  
 A;Note: the source is designated as *Halobacterium marismortui*  
 A;Note: the protein is designated as ribosomal protein S6  
 A;Note: the protein was extracted from the 30S small ribosomal subunit  
 C;Superfamily: rat ribosomal protein L7a  
 C;Keywords: protein biosynthesis; ribosome

Query Match 72.5%; Score 29; DB 1; Length 116;  
 Best Local Similarity 75.0%; Pred. No. 42;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 :||| |||  
 Db 46 ELVFVAED 53

RESULT 36

T06645

hypothetical protein T20K18.220 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Jun-2000

C;Accession: T06645

R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, April 1999

A;Reference number: Z15790

A;Accession: T06645

A;Molecule type: DNA

A;Residues: 1-152 <BEV>

A;Cross-references: EMBL:AL049640; GSPDB:GN00062; ATSP:T20K18.220

A;Experimental source: cultivar Columbia; BAC clone T20K18

C;Genetics:

A;Gene: ATSP:T20K18.220

A;Map position: 4

A;Introns: 87/3; 109/3

C;Superfamily: Arabidopsis thaliana hypothetical protein F7A7.100

Query Match 72.5%; Score 29; DB 2; Length 152;

Best Local Similarity 100.0%; Pred. No. 56;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFA 6  
 |||||  
 Db 9 KLVFFA 14

RESULT 37

T13487

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Monochoria vaginalis chloroplast (fragment)

C;Species: chloroplast Monochoria vaginalis

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 03-Jun-2002

C;Accession: T13487

R;Graham, S.W.; Kohn, J.R.; Morton, B.R.; Eckenwalder, J.E.; Barrett, S.C. Syst. Biol. 47, 545-567, 1998

A;Title: Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae.

A;Reference number: Z17600

A;Accession: T13487

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-162 <GRA>

A;Cross-references: EMBL:U41616; NID:g1173991; PID:g1173992; PIDN:AAD09707.1

C;Genetics:

A;Genome: chloroplast

A;Note: ndhF

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5

C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match 72.5%; Score 29; DB 2; Length 162;

Best Local Similarity 50.0%; Pred. No. 59;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 38

T13563

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - *Pontederia cordata* chloroplast (fragment)

C;Species: chloroplast *Pontederia cordata*

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 03-Jun-2002

C;Accession: T13563

R;Graham, S.W.; Kohn, J.R.; Morton, B.R.; Eckenwalder, J.E.; Barrett, S.C. Syst. Biol. 47, 545-567, 1998

A;Title: Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae.

A;Reference number: Z17600

A;Accession: T13563

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-162 <GRA>

A;Cross-references: EMBL:U41619; NID:g1174003; PID:g1174004; PIDN:AAD09710.1

A;Experimental source: var. ovalis

C;Genetics:

A;Genome: chloroplast

A;Note: ndhF

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5

C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match 72.5%; Score 29; DB 2; Length 162;

Best Local Similarity 50.0%; Pred. No. 59;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 39

T13656

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - *Pontederia rotundifolia* chloroplast (fragment)

C;Species: chloroplast *Pontederia rotundifolia*

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 03-Jun-2002

C;Accession: T13656

R;Graham, S.W.; Kohn, J.R.; Morton, B.R.; Eckenwalder, J.E.; Barrett, S.C. Syst. Biol. 47, 545-567, 1998

A;Title: Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae.

A;Reference number: Z17600

A;Accession: T13656

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-162 <GRA>

A;Cross-references: EMBL:U41620; NID:g1174011; PID:g1174012; PIDN:AAD09711.1



C;Genetics:  
A;Genome: chloroplast  
A;Note: ndhF  
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5  
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match 72.5%; Score 29; DB 2; Length 162;  
Best Local Similarity 50.0%; Pred. No. 59;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 40

T13659

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - *Pontederia sagittata* chloroplast (fragment)

C;Species: chloroplast *Pontederia sagittata*

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 03-Jun-2002

C;Accession: T13659

R;Graham, S.W.; Kohn, J.R.; Morton, B.R.; Eckenwalder, J.E.; Barrett, S.C. Syst. Biol. 47, 545-567, 1998

A;Title: Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae.

A;Reference number: Z17600

A;Accession: T13659

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-162 <GRA>

A;Cross-references: EMBL:U41621; NID:g1174015; PID:g1174016; PIDN:AAD09712.1

C;Genetics:

A;Genome: chloroplast

A;Note: ndhF

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5

C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match 72.5%; Score 29; DB 2; Length 162;  
Best Local Similarity 50.0%; Pred. No. 59;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 41

T13562

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - *Pontederia cordata* chloroplast (fragment)

C;Species: chloroplast *Pontederia cordata*

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 03-Jun-2002

C;Accession: T13562

R;Graham, S.W.; Kohn, J.R.; Morton, B.R.; Eckenwalder, J.E.; Barrett, S.C. Syst. Biol. 47, 545-567, 1998

A;Title: Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae.  
 A;Reference number: Z17600  
 A;Accession: T13562  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-164 <GRA>  
 A;Cross-references: EMBL:U41617; NID:g1173999; PID:g1174000; PIDN:AAD09708.1  
 A;Experimental source: var. codata  
 C;Genetics:  
 A;Genome: chloroplast  
 A;Note: ndhF  
 C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5  
 C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Query Match 72.5%; Score 29; DB 2; Length 164;  
 Best Local Similarity 50.0%; Pred. No. 60;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |::||::|  
 Db 13 KVIFFSKD 20

#### RESULT 42

S39864

late competence operon required for DNA binding and uptake comEB - *Bacillus subtilis*

C;Species: *Bacillus subtilis*

C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Jun-2000

C;Accession: S39864; D69602

R;Hahn, J.; Inamine, G.; Kozlov, Y.; Dubnau, D.

Mol. Microbiol. 10, 99-111, 1993

A;Title: Characterization of comE, a late competence operon of *Bacillus subtilis* required for the binding and uptake of transforming DNA.

A;Reference number: S39862; MUID:95058187; PMID:7968523

A;Accession: S39864

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-189 <HAH>

A;Cross-references: EMBL:L15202; NID:g289258; PIDN:AAC36906.1; PID:g289261

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.;

Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;

Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.;

Connerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.;

Duesterhoeft, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;

Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.;

Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;

Guiseppe, G.; Guy, B.J.; Haga, K.; Haiech, J.; Harwood, C.R.; Henaut, A.;

Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.;

Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.;

Kobayashi, Y.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.;

Lapidus, A.; Lardinois, S.

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwollik, S.; Prescott, A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.; Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenegger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: D69602

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-189 <KUN>

A;Cross-references: GB:Z99117; GB:AL009126; NID:g2634966; PIDN:CAB14500.1; PID:g2635004

A;Experimental source: strain 168

C;Genetics:

A;Gene: comEB

A;Start codon: GTG

C;Superfamily: Archaeoglobus probable dCMP deaminase

Query Match 72.5%; Score 29; DB 2; Length 189;

Best Local Similarity 62.5%; Pred. No. 70;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| |::|||  
Db 111 KTVYYAED 118

RESULT 43

H85138

hypothetical protein AT4g12900 [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001

C;Accession: H85138

R;anonymous, The European Union *Arabidopsis* Genome Sequencing Consortium, The Cold Spring Harbor, Washington University in St Louis and PE Biosystems *Arabidopsis* Sequencing Consortium.

Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: H85138

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-231 <STO>

A;Cross-references: GB:NC\_001268; NID:g7267992; PIDN:CAB78332.1; GSPDB:GN00140

C;Genetics:  
A;Gene: AT4g12900  
A;Map position: 4  
C;Superfamily: Arabidopsis thaliana hypothetical protein F7A7.100

Query Match 72.5%; Score 29; DB 2; Length 231;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFA 6  
| | | | |  
Db 12 KLVFFA 17

RESULT 44

B86301

hypothetical protein F19K19.10 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C;Accession: B86301

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;  
Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.;  
Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.;  
Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.;  
Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.;  
Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.;  
Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-  
Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu,  
S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda,  
M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.;  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.;  
Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken,  
S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.;  
Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B86301

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-247 <STO>

A;Cross-references: GB:AE005172; NID:g9989059; PIDN:AAG10822.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

Query Match 72.5%; Score 29; DB 2; Length 247;  
Best Local Similarity 71.4%; Pred. No. 91;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
: | | | : | |  
Db 169 VVFFSED 175

RESULT 45

S39747

ywfN protein - *Bacillus subtilis*

N;Alternate names: hypothetical protein ipa-92r

C;Species: *Bacillus subtilis*

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

C;Accession: S39747; F70056

R;Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu, M.; Lubochinsky, B.; Marcelino, L.; Moszer, I.; Presecan, E.; Santana, M.; Schneider, E.; Schweizer, J.; Vertes, A.; Rapoport, G.; Danchin, A.  
Mol. Microbiol. 10, 371-384, 1993

A;Title: *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region from 325 degrees to 333 degrees.

A;Reference number: S39655; MUID:95020537; PMID:7934828

A;Accession: S39747

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-258 <GLA>

A;Cross-references: EMBL:X73124; NID:g413923; PIDN:CAA51648.1; PID:g414016

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Guiseppe, G.; Guy, B.J.; Haga, K.; Haiech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwollik, S.; Prescott, A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.; Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenegger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: F70056

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-258 <KUN>

A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15789.1;  
PID:g2636298

A;Experimental source: strain 168

C;Genetics:

A;Gene: ywfN

C;Superfamily: hypothetical protein ylbO

Query Match 72.5%; Score 29; DB 1; Length 258;  
Best Local Similarity 75.0%; Pred. No. 95;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||| | ||  
Db 231 KLVLFEEED 238

#### RESULT 46

T06661

hypothetical protein T6G15.140 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 22-Oct-1999

C;Accession: T06661

R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.;  
Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, April 1999

A;Reference number: Z15791

A;Accession: T06661

A;Molecule type: DNA

A;Residues: 1-273 <BEV> .

A;Cross-references: EMBL:AL049656; GSPDB:GN00062; ATSP:T6G15.140

A;Experimental source: cultivar Columbia; BAC clone T6G15

C;Genetics:

A;Gene: ATSP:T6G15.140

A;Map position: 4

A;Introns: 48/1; 95/3; 111/3; 144/1; 220/3; 237/3

Query Match 72.5%; Score 29; DB 2; Length 273;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAE 7  
|||||  
Db 216 LVFFAE 221

#### RESULT 47

H64054

site-specific DNA-methyltransferase (adenine-specific) (EC 2.1.1.72) dam -  
Haemophilus influenzae (strain Rd KW20)

C;Species: Haemophilus influenzae

C;Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 05-May-2000

C;Accession: H64054

R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.;  
Kerlavage, A.R.; Bult, C.J.; Tomb, J.F.; Dougherty, B.A.; Merrick, J.M.;

McKenney, K.; Sutton, G.; FitzHugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.;

Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.F.; Phillips, C.A.;

Spriggs, T.; Hedblom, E.; Cotton, M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen,

D.T.; Saudek, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
 Science 269, 496-512, 1995  
 A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.  
 A;Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.  
 A;Reference number: A64000; MUID:95350630; PMID:7542800  
 A;Accession: H64054  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-286 <TIGR>  
 A;Cross-references: GB:U32705; GB:L42023; NID:g1573156; PIDN:AAC21877.1; PID:g1573168; TIGR:HI0209  
 C;Superfamily: site-specific methyltransferase (adenine-specific) EcoRV  
 C;Keywords: methyltransferase; S-adenosylmethionine

Query Match 72.5%; Score 29; DB 2; Length 286;  
 Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 | :|||:|  
 Db 90 KPIFFADD 97

#### RESULT 48

AC1342

ABC transporter (ATP-binding protein) homolog lmo2139 [imported] - *Listeria monocytogenes* (strain EGD-e)

C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C;Accession: AC1342

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AC1342

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-300 <GLA>

A;Cross-references: GB:NC\_003210; PIDN:CAD00217.1; PID:g16411609; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo2139

C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 72.5%; Score 29; DB 2; Length 300;  
Best Local Similarity 71.4%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
:|:||||  
Db 92 QLIFFAE 98

RESULT 49

AI1712

ABC transporter (ATP-binding protein) homolog lin2244 [imported] - *Listeria innocua* (strain Clip11262)

C;Species: *Listeria innocua*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C;Accession: AI1712

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AI1712

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-300 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC97473.1; PID:g16414757; GSPDB:GN00178

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: lin2244

C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 72.5%; Score 29; DB 2; Length 300;  
Best Local Similarity 71.4%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
:|:||||  
Db 92 QLIFFAE 98

RESULT 50

F86805

cation transporter yogJ [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)

C;Species: *Lactococcus lactis* subsp. *lactis*

C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001

C;Accession: F86805



R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich, S.D.; Sorokin, A.  
 Genome Res. 11, 731-753, 2001  
 A;Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403.  
 A;Reference number: A86625; MUID:21235186; PMID:11337471  
 A;Accession: F86805  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-314 <STO>  
 A;Cross-references: GB:AE005176; PID:g12724436; PIDN:AAK05544.1; GSPDB:GN00146  
 A;Experimental source: strain IL1403  
 C;Genetics:  
 A;Gene: yogJ

Query Match 72.5%; Score 29; DB 2; Length 314;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
 || ||||  
 Db 163 LVIFAED 169

# RESULT 51

AH2097

sorbitol dehydrogenase [imported] - *Nostoc* sp. (strain PCC 7120)

C;Species: *Nostoc* sp. PCC 7120

A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C;Accession: AH2097

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium* *Anabaena* sp. strain PCC 7120.

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AH2097

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-352 <KUR>

A;Cross-references: GB:BA000019; PIDN:BA074034.1; PID:g17131427; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr2335

C;Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

Query Match 72.5%; Score 29; DB 2; Length 352;  
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
 |::|||  
 Db 262 KILFFAE 268

RESULT 52

T48903

wax synthase [imported] - *Simmondsia chinensis*

C;Species: *Simmondsia chinensis*

C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C;Accession: T48903

R;Lardizabal, K.D.; Metz, J.G.; Sakamoto, T.; Hutton, W.C.; Pollard, M.R.; Lassner, M.W.

Plant Physiol. 122, 645-655, 2000

A;Title: Purification of a jojoba embryo wax synthase, cloning of its cDNA and production of high levels of wax in seeds of transgenic *Arabidopsis*.

A;Reference number: Z25002

A;Accession: T48903

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-352 <LAR>

A;Cross-references: EMBL:AF149919; PIDN:AAD38041.1

Query Match 72.5%; Score 29; DB 2; Length 352;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFA 6

|||||

Db 135 KLVFFA 140

RESULT 53

E97128

magnesium and cobalt transport protein CAC1852 [imported] - *Clostridium acetobutylicum*

C;Species: *Clostridium acetobutylicum*

C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 01-Mar-2002

C;Accession: E97128

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, H.M.; Dubois, J.; Qiu, D.; Hitti, J.; Wolf, Y.I.; Tatusov, R.L.; Sabathe, F.; Doucette-Stamm, L.; Soucaille, P.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum*.

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: E97128

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-354 <KUR>

A;Cross-references: GB:AE001437; PIDN:AAK79816.1; PID:g15024829; GSPDB:GN00168

A;Experimental source: *Clostridium acetobutylicum* ATCC824

C;Genetics:

A;Gene: CAC1852

C;Superfamily: magnesium and cobalt transport protein

Query Match 72.5%; Score 29; DB 2; Length 354;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy           1 KLVFFAED 8  
             |:::| ||  
Db           114 KMIYFRED 121

RESULT 54

S59678

HST2 protein - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: protein LPA2c; protein YPL015c

C;Species: *Saccharomyces cerevisiae*

C;Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 29-Oct-1999

C;Accession: S59678

R;Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; Wang, Y.; Winnett, E.

submitted to the EMBL Data Library, August 1995

A;Description: The sequence of *Saccharomyces cerevisiae* chromosome XVI left arm.

A;Reference number: S59677

A;Accession: S59678

A;Molecule type: DNA

A;Residues: 1-357 <HAL>

A;Cross-references: EMBL:U33335; NID:g965076; PIDN:AAB68090.1; PID:g965078;

MIPS:YPL015c

C;Genetics:

A;Gene: SGD:HST2

A;Cross-references: SGD:S0005936; MIPS:YPL015c

A;Map position: 16L

C;Keywords: transmembrane protein

F;218-234/Domain: transmembrane #status predicted <TMM>

Query Match                   72.5%;   Score 29;   DB 2;   Length 357;  
Best Local Similarity   71.4%;   Pred. No. 1.3e+02;  
Matches       5;   Conservative       1;   Mismatches       1;   Indels       0;   Gaps       0;

Qy           2 LVFFAED 8  
             :| | | ||  
Db           181 IVFFGED 187

RESULT 55

A75091

phosphonoacetate hydrolase related PAB2429 - *Pyrococcus abyssi* (strain Orsay)

C;Species: *Pyrococcus abyssi*

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999

C;Accession: A75091

R;anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure and evolution.

A;Reference number: A75001

A;Accession: A75091

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-369 <KAW>

A;Cross-references: GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB50030.1;

PID:e1515927; PID:g5458542

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB2429

Query Match 72.5%; Score 29; DB 2; Length 369;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| ||| ||  
Db 269 KSVFFRED 276

RESULT 56

T48008

hypothetical protein T17J13.120 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C;Accession: T48008

R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.;  
Mayer, K.F.X.; Quetier, F.; Salanoubat, M.

submitted to the Protein Sequence Database, February 2000

A;Reference number: Z24482

A;Accession: T48008

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-428 <RIE>

A;Cross-references: EMBL:AL138651

A;Experimental source: cultivar Columbia; BAC clone T17J13

C;Genetics:

A;Map position: 3

A;Introns: 137/3

A;Note: T17J13.120

Query Match 72.5%; Score 29; DB 2; Length 428;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFA 6  
|||||  
Db 271 KLVFFA 276

RESULT 57

A64422

conserved hypothetical protein MJ0977 - Methanococcus jannaschii

C;Species: Methanococcus jannaschii

C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C;Accession: A64422

R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.;  
Blake, J.A.; FitzGerald, L.M.; Clayton, R.A.; Gocayne, J.D.; Kerlavage, A.R.;  
Dougherty, B.A.; Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.; Kirkness,  
E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Scott, J.L.; Geoghagen,  
N.S.M.; Weidman, J.F.; Fuhrmann, J.L.; Nguyen, D.; Utterback, T.R.; Kelley,  
J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.;  
Hurst, M.A.

Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.;  
Woese, C.R.; Venter, J.C.

A;Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*.  
A;Reference number: A64300; MUID:96337999; PMID:8688087  
A;Accession: A64422  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-469 <BUL>  
A;Cross-references: GB:U67540; GB:L77117; NID:g1591631; PIDN:AAB98980.1;  
PID:g1591640; TIGR:MJ0977  
C;Genetics:  
A;Map position: FOR909596-911005  
C;Superfamily: *Methanococcus jannaschii* conserved hypothetical protein MJ0977

Query Match                    72.5%;   Score 29;   DB 2;   Length 469;  
Best Local Similarity       75.0%;   Pred. No. 1.8e+02;  
Matches       6;   Conservative       1;   Mismatches       1;   Indels       0;   Gaps       0;

Qy                1 KLVFFAED 8  
                  ||: ||||  
Db                417 KLLCFAED 424

RESULT 58

T36945  
hypothetical protein SCJ1.12 - *Streptomyces coelicolor*  
C;Species: *Streptomyces coelicolor*  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C;Accession: T36945  
R;Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.;  
Rajandream, M.A.  
submitted to the EMBL Data Library, September 1999  
A;Reference number: Z21607  
A;Accession: T36945  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-508 <SEE>  
A;Cross-references: EMBL:AL109962; PIDN:CAB53130.1; GSPDB:GN00070;  
SCOEDB:SCJ1.12  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCOEDB:SCJ1.12

Query Match                    72.5%;   Score 29;   DB 2;   Length 508;  
Best Local Similarity       71.4%;   Pred. No. 1.9e+02;  
Matches       5;   Conservative       1;   Mismatches       1;   Indels       0;   Gaps       0;

Qy                2 LVFFAED 8  
                  :||| ||  
Db                25 MVFFVED 31

RESULT 59

T45820  
hypothetical protein F2K15.10 - *Arabidopsis thaliana*  
N;Alternate names: hypothetical protein T2J13.10  
C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 02-Jun-2000

C;Accession: T45820; T46115  
 R;Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.;  
 Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.  
 submitted to the Protein Sequence Database, January 2000  
 A;Reference number: Z23015  
 A;Accession: T45820  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-590 <RIE>  
 A;Cross-references: EMBL:AL132956  
 A;Experimental source: cultivar Columbia; BAC clone F2K15  
 R;Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.;  
 Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetier, F.; Salanoubat, M.  
 submitted to the Protein Sequence Database, November 1999  
 A;Reference number: Z23023  
 A;Accession: T46115  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-165 <RI2>  
 A;Cross-references: EMBL:AL132967  
 A;Experimental source: cultivar Columbia; BAC clone T2J13  
 C;Genetics:  
 A;Map position: 3  
 A;Introns: 335/3; 361/3; 450/2  
 A;Note: F2K15.10; T2J13.10

Query Match 72.5%; Score 29; DB 2; Length 590;  
 Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 :| || ||  
 Db 223 RLTFECED 230

# RESULT 60

T24172  
 hypothetical protein R11A8.4 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T24172  
 R;Cummings, P.  
 submitted to the EMBL Data Library, March 1996  
 A;Reference number: Z19849  
 A;Accession: T24172  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-607 <WIL>  
 A;Cross-references: EMBL:Z70310; PIDN:CAA94364.1; GSPDB:GN00022; CESP:R11A8.4  
 A;Experimental source: clone R11A8  
 C;Genetics:  
 A;Gene: CESP:R11A8.4  
 A;Map position: 4  
 A;Introns: 18/3; 68/3; 111/3; 206/3; 500/3; 565/3

Query Match 72.5%; Score 29; DB 2; Length 607;  
 Best Local Similarity 71.4%; Pred. No. 2.3e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
: ||| ||  
Db 298 IVFFGED 304

RESULT 61

A82153

PTS system, fructose-specific IIABC component VC1822 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)

C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 01-Mar-2002

C;Accession: A82153

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleishmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: A82153

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-621 <HEI>

A;Cross-references: GB:AE004258; GB:AE003852; NID:g9656343; PIDN:AAF94970.1; GSPDB:GN00126; TIGR:VC1822

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC1822

A;Map position: 1

C;Superfamily: phosphotransferase system enzyme II, fructose-specific; phosphotransferase system mannitol-specific enzyme II factor III homology

Query Match 72.5%; Score 29; DB 2; Length 621;  
Best Local Similarity 75.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | |  
Db 580 KLVIFAHD 587

RESULT 62

S61435

metalloproteinase (EC 3.4.24.-) precursor - *Aspergillus fumigatus*

C;Species: *Aspergillus fumigatus*

C;Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 29-Jan-1999

C;Accession: S61435; S61436; S61450

R;Jaton-Ogay, K.; Paris, S.; Huerre, M.; Quadroni, M.; Falchetto, R.; Togni, G.; Latge, J.P.; Monod, M.

Mol. Microbiol. 14, 917-928, 1994

A;Title: Cloning and disruption of the gene encoding an extracellular metalloprotease of *Aspergillus fumigatus*.

A;Reference number: S61435; MUID:95231298; PMID:7715453  
 A;Accession: S61435  
 A;Molecule type: DNA  
 A;Residues: 1-633 <JAT>  
 A;Cross-references: EMBL:Z30424; NID:g458475; PID:g458476  
 A;Experimental source: strain delta18  
 A;Accession: S61436  
 A;Molecule type: mRNA  
 A;Residues: 1-633 <JAW>  
 A;Cross-references: EMBL:Z30424  
 A;Accession: S61450  
 A;Molecule type: protein  
 A;Residues: 246-276;362-371;407-416;419-428;477-488;571-588 <JAF>  
 C;Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F;1-18/Domain: signal sequence #status predicted <SIG>  
 F;19-245/Domain: propeptide #status predicted <PRO>  
 F;246-633/Product: metalloproteinase #status experimental <MAT>  
 F;428,432/Binding site: zinc, catalytic (His) (active) #status predicted  
 F;429/Active site: Glu #status predicted

Query Match 72.5%; Score 29; DB 2; Length 633;  
 Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||:| :|  
 Db 198 KLVYFVKD 205

# RESULT 63

S42894  
 metalloproteinase MEP - *Aspergillus fumigatus*  
 C;Species: *Aspergillus fumigatus*  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Sep-1997  
 C;Accession: S42894  
 R;Jaton-Ogay, K.; Quadroni, M.; Falchetto, R.; Togni, G.; Paris, S.; Latge, J.;  
 Monod, M.  
 submitted to the EMBL Data Library, March 1994  
 A;Description: Nucleotide sequence of genomic and cDNA clones encoding an  
 extracellular metalloprotease of *Aspergillus fumigatus*.  
 A;Reference number: S42894  
 A;Accession: S42894  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-649 <JAT>  
 A;Cross-references: EMBL:Z30424; NID:g458475; PID:g458476  
 C;Genetics:  
 A;Introns: 119/3; 251/2; 588/2

Query Match 72.5%; Score 29; DB 2; Length 649;  
 Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||:| :|  
 Db 198 KLVYFVKD 205



RESULT 64

A96540

hypothetical protein F14I3.4 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 03-Jun-2002

C;Accession: A96540

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A96540

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-710 <STO>

A;Cross-references: GB:AE005173; NID:g5734786; PIDN:AAD50051.1; GSPDB:GN00141

C;Genetics:

A;Gene: F14I3.4

A;Map position: 1

C;Superfamily: prolyl oligopeptidase

Query Match 72.5%; Score 29; DB 2; Length 710;

Best Local Similarity 75.0%; Pred. No. 2.7e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8

||| :|||

Db 158 KLVAYAED 165

RESULT 65

JH0595

potassium channel protein cdrK - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 20-Aug-1999

C;Accession: JH0595

R;Hwang, P.M.; Glatt, C.E.; Bredt, D.S.; Yellen, G.; Snyder, S.H.  
Neuron 8, 473-481, 1992

A;Title: A novel K+ channel with unique localizations in mammalian brain: molecular cloning and characterization.

A;Reference number: JH0595; MUID:92198655; PMID:1550672

A;Accession: JH0595

A;Molecule type: mRNA  
 A;Residues: 1-802 <HWA>  
 A;Cross-references: GB:M77482; NID:g203395; PIDN:AAA40905.1; PID:g203396  
 A;Experimental source: circumvallate papillae  
 C;Superfamily: potassium channel protein drkl  
 C;Keywords: glycoprotein; phosphoprotein; transmembrane protein  
 F;191-212/Domain: transmembrane #status predicted <TM1>  
 F;233-254/Domain: transmembrane #status predicted <TM2>  
 F;265-286/Domain: transmembrane #status predicted <TM3>  
 F;299-320/Domain: transmembrane #status predicted <TM4>  
 F;335-356/Domain: transmembrane #status predicted <TM5>  
 F;397-418/Domain: transmembrane #status predicted <TM6>  
 F;187,287,419,446/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;448,500/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 72.5%; Score 29; DB 2; Length 802;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAE 7  
 |||||  
 Db 352 LVFFAE 357

#### RESULT 66

##### CHRTD1

potassium channel protein drkl - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 22-Jun-1999

C;Accession: S05448; A44838

R;Frech, G.C.; VanDongen, A.M.J.; Schuster, G.; Brown, A.M.; Joho, R.H.

Nature 340, 642-645, 1989

A;Title: A novel potassium channel with delayed rectifier properties isolated from rat brain by expression cloning.

A;Reference number: S05448; MUID:89365157; PMID:2770868

A;Accession: S05448

A;Molecule type: mRNA

A;Residues: 1-853 <FRE>

A;Cross-references: EMBL:X16476; NID:g57785; PIDN:CAA34497.1; PID:g57786

A;Note: it is uncertain whether Met-1 or Met-17 is the initiator

R;Drewe, J.A.; Verma, S.; Frech, G.; Joho, R.H.

J. Neurosci. 12, 538-548, 1992

A;Title: Distinct spatial and temporal expression patterns of K<sup>+</sup> channel mRNAs from different subfamilies.

A;Reference number: A44838; MUID:92156897; PMID:1740690

A;Accession: A44838

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 'MPAG',1-571 <DRE>

A;Cross-references: GB:M81783; NID:g205038

A;Experimental source: brain

A;Note: sequence extracted from NCBI backbone (NCBIP:81768)

C;Genetics:

A;Gene: drkl

C;Superfamily: potassium channel protein drkl

C;Keywords: glycoprotein; ion channel; phosphoprotein; potassium channel; transmembrane protein

F;1-182/Domain: intracellular #status predicted <INT1>  
 F;183-204/Domain: transmembrane #status predicted <TM1>  
 F;225-245/Domain: transmembrane #status predicted <TM2>  
 F;256-276/Domain: transmembrane #status predicted <TM3>  
 F;291-312/Domain: transmembrane #status predicted <TM4>  
 F;327-348/Domain: transmembrane #status predicted <TM5>  
 F;389-410/Domain: transmembrane #status predicted <TM6>  
 F;411-853/Domain: intracellular #status predicted <INT2>  
 F;279/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 72.5%; Score 29; DB 1; Length 853;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAE 7  
 |||||  
 Db 344 LVFFAE 349

# RESULT 67

I56529

potassium channel protein - mouse

C;Species: Mus musculus (house mouse)

C;Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Aug-1999

C;Accession: I56529

R;Pak, M.D.; Covarrubias, M.; Ratcliffe, A.; Salkoff, L.

J. Neurosci. 11, 869-880, 1991

A;Title: A mouse brain homolog of the Drosophila Shab K+ channel with conserved delayed-rectifier properties.

A;Reference number: I56529; MUID:91162315; PMID:2002364

A;Accession: I56529

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-857 <RES>

A;Cross-references: GB:M64228; NID:g200975; PIDN:AAA40112.1; PID:g200976

C;Genetics:

A;Gene: Shab

C;Superfamily: potassium channel protein drk1

Query Match 72.5%; Score 29; DB 2; Length 857;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAE 7  
 |||||  
 Db 348 LVFFAE 353

# RESULT 68

S31761

potassium channel protein DRK1 - human

C;Species: Homo sapiens (man)

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Aug-1999

C;Accession: S31761

R;Albrecht, B.; Lorra, C.; Stocker, K.; Pongs, O.

submitted to the EMBL Data Library, September 1992

A;Description: Cloning, expression and chromosomal localization of the delayed rectifier type K<sup>+</sup> channel human 2.1 (h-DRK1) gene.

A;Reference number: S31761

A;Accession: S31761

A;Molecule type: DNA

A;Residues: 1-858 <ALB>

A;Cross-references: EMBL:X68302; NID:g30892; PIDN:CAA48374.1; PID:g30893

C;Genetics:

A;Gene: GDB:KCNB1; KV2.1; DRK1

A;Cross-references: GDB:128081; OMIM:600397

A;Map position: 20q13.2-20q13.2

C;Superfamily: potassium channel protein drk1

Query Match                      72.5%;   Score 29;   DB 2;   Length 858;  
Best Local Similarity      100.0%;   Pred. No. 3.2e+02;  
Matches       6;   Conservative       0;   Mismatches       0;   Indels       0;   Gaps       0;

QY                      2 LVFFAE 7  
                         |||||  
Db                      348 LVFFAE 353

#### RESULT 69

S62752

probable DNA-directed DNA polymerase (EC 2.7.7.7) - Gelasinospora sp.

mitochondrion plasmid kal

C;Species: mitochondrion Gelasinospora sp.

C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 07-Dec-1999

C;Accession: S62752; S72369

R;Yuewang, W.; Yang, X.; Griffiths, A.J.F.

Curr. Genet. 29, 150-158, 1996

A;Title: Structure of a Gelasinospora linear plasmid closely related to the kalilo plasmid of Neurospora intermedia.

A;Reference number: S62751; MUID:96418877; PMID:8821662

A;Accession: S62752

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-987 <YUE>

A;Cross-references: EMBL:L40494

A;Experimental source: strain G114

R;Yuewang, W.; Yang, X.; Griffiths, A.J.F.

submitted to the EMBL Data Library, October 1995

A;Description: Structure of a Gelasinospora linear plasmid closely related to the kalilo plasmid of Neurospora intermedia.

A;Reference number: S72369

A;Accession: S72369

A;Molecule type: DNA

A;Residues: 1-216,'LS',219-987 <YUW>

A;Cross-references: EMBL:L40494; NID:g972729; PIDN:AAB41447.1; PID:g1805212

C;Genetics:

A;Genome: mitochondrion plasmid kal

A;Genetic code: SGC3

C;Superfamily: Neurospora crassa mitochondrion plasmid probable DNA-directed DNA polymerase

C;Keywords: mitochondrion; nucleotidyltransferase

Query Match                      72.5%;   Score 29;   DB 2;   Length 987;

Best Local Similarity 62.5%; Pred. No. 3.7e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|:|||||:  
Db 327 KLVFFTEN 334

RESULT 70

S30236

genome polyprotein - zucchini yellow mosaic virus (strain Singapore) (fragment)  
N;Contains: coat protein; nuclear inclusion protein a; RNA-directed RNA  
polymerase (EC 2.7.7.48); VPg protein

C;Species: zucchini yellow mosaic virus, ZYMV

A;Variety: strain Singapore

C;Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 17-Nov-2000

C;Accession: S30236

R;Wu, M.; Yeong, C.Y.; Lee, S.C.; Wong, S.M.

Nucleic Acids Res. 21, 1317, 1993

A;Title: Nucleotide sequence of the 3' half of zucchini yellow mosaic virus  
(Singapore isolate) genome encoding the 4K protein, protease, polymerase and  
coat protein.

A;Reference number: S30236; MUID:93219099; PMID:8464715

A;Accession: S30236

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: mRNA

A;Residues: 1-1016 <WUM>

A;Cross-references: EMBL:X68509; NID:g288233; PIDN:CAA48521.1; PID:g938312

A;Note: the nucleotide sequence was submitted to the EMBL Data Library,  
September 1992

C;Superfamily: tobacco etch virus genome polyprotein

C;Keywords: coat protein; genome-linked protein; nucleotidyltransferase;  
phosphoprotein; polyprotein

F;9-61/Product: VPg protein #status predicted <VPG>

F;62-494/Product: nuclear inclusion protein a #status predicted <NIA>

F;495-1010/Product: RNA-directed RNA polymerase #status predicted <POL>

F;1011-1016/Product: coat protein (fragment) #status predicted <COP>

F;125/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Query Match 72.5%; Score 29; DB 2; Length 1016;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFA 6  
|||||  
Db 837 KLVFFA 842

RESULT 71

T34418

hypothetical protein F12F3.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C;Accession: T34418

R;Fulton, B.; Wohldmann, P.

submitted to the EMBL Data Library, July 1998

A;Description: The sequence of C. elegans cosmid F12F3.

A;Reference number: Z21521  
A;Accession: T34418  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-3488 <FUL>  
A;Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3  
A;Experimental source: strain Bristol N2; clone F12F3  
C;Genetics:  
A;Gene: CESP:F12F3.3  
A;Map position: 5  
A;Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 72.5%; Score 29; DB 2; Length 3488;  
Best Local Similarity 75.0%; Pred. No. 1.4e+03;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|| | |||  
Db 2678 KLTFVAED 2685

#### RESULT 72

S17231

dynein beta heavy chain, ciliary - sea urchin (*Anthocidaris crassispina*)

N;Contains: dynein ATPase (EC 3.6.4.2)

C;Species: *Anthocidaris crassispina*

C;Date: 30-Sep-1991 #sequence\_revision 02-May-1994 #text\_change 19-Apr-2002

C;Accession: S17231; PS0415

R;Ogawa, K.

Nature 352, 643-645, 1991

A;Title: Four ATP-binding sites in the midregion of the beta heavy chain of dynein.

A;Reference number: S17231; MUID:91326104; PMID:1830928

A;Accession: S17231

A;Molecule type: mRNA

A;Residues: 1-4466 <OGA>

A;Cross-references: GB:D01021; NID:g217202; PIDN:BAA00827.1; PID:g217203

R;Ogawa, K.

Proc. Jpn. Acad. B Phys. Biol. Sci. 67, 27-31, 1991

A;Title: ATP-binding site in dynein beta-heavy chain: identification by molecular cloning.

A;Reference number: PS0415

A;Accession: PS0415

A;Molecule type: mRNA

A;Residues: 764-1001,'APQ',1005-2036,'VPSSVET' <OG2>

C;Superfamily: dynein heavy chain, ciliary

C;Keywords: ATP; heterotetramer; hydrolase; microtubule binding; nucleotide binding; P-loop

F;154-161/Region: nucleotide-binding motif A (P-loop)

F;1852-1859/Region: nucleotide-binding motif A (P-loop)

F;2133-2140/Region: nucleotide-binding motif A (P-loop)

F;2460-2467/Region: nucleotide-binding motif A (P-loop)

F;2805-2812/Region: nucleotide-binding motif A (P-loop)

F;160/Binding site: ATP (Lys) #status predicted

F;1858/Binding site: ATP (Lys) #status predicted

F;2139/Binding site: ATP (Lys) #status predicted

F;2466/Binding site: ATP (Lys) #status predicted

F;2811/Binding site: ATP (Lys) #status predicted

Query Match                    72.5%;    Score 29;    DB 1;    Length 4466;  
Best Local Similarity    62.5%;    Pred. No. 1.7e+03;  
Matches       5;    Conservative       2;    Mismatches       1;    Indels       0;    Gaps       0;  
  
Qy                1 KLVFFAED 8  
                  |||:| :|  
Db               2520 KLVYFIDD 2527

RESULT 73

S17653

dynein beta heavy chain, ciliary - sea urchin (Tripneustes gratilla)

N;Contains: dynein ATPase (EC 3.6.4.2)

C;Species: Tripneustes gratilla

C;Date: 04-Dec-1992 #sequence\_revision 02-May-1994 #text\_change 19-Apr-2002

C;Accession: S17653; S24628

R;Gibbons, I.R.; Gibbons, B.H.; Mocz, G.; Asai, D.J.

Nature 352, 640-643, 1991

A;Title: Multiple nucleotide-binding sites in the sequence of dynein beta heavy chain.

A;Reference number: S17653; MUID:91326103; PMID:1830927

A;Accession: S17653

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-4466 <GIB1>

A;Cross-references: EMBL:X59603; NID:g10709; PIDN:CAA42170.1; PID:g10710

A;Accession: S24628

A;Molecule type: protein

A;Residues: 162-172;1193-1204;3240-3259;3325-3339 <GIB2>

C;Superfamily: dynein heavy chain, ciliary

C;Keywords: ATP; heterotetramer; hydrolase; microtubule binding; nucleotide binding; P-loop

F;154-161/Region: nucleotide-binding motif A (P-loop)

F;1852-1859/Region: nucleotide-binding motif A (P-loop)

F;2133-2140/Region: nucleotide-binding motif A (P-loop)

F;2460-2467/Region: nucleotide-binding motif A (P-loop)

F;2805-2812/Region: nucleotide-binding motif A (P-loop)

F;1858/Binding site: ATP (Lys) #status predicted

F;2139/Binding site: ATP (Lys) #status predicted

F;2466/Binding site: ATP (Lys) #status predicted

F;2811/Binding site: ATP (Lys) #status predicted

Query Match                    72.5%;    Score 29;    DB 1;    Length 4466;  
Best Local Similarity    62.5%;    Pred. No. 1.7e+03;  
Matches       5;    Conservative       2;    Mismatches       1;    Indels       0;    Gaps       0;  
  
Qy                1 KLVFFAED 8  
                  |||:| :|  
Db               2520 KLVYFIDD 2527

RESULT 74

B44358

cysteine-rich protein hCRP homolog - chicken (fragment)

C;Species: Gallus gallus (chicken)

C;Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Jun-2001  
C;Accession: B44358  
R;Sadler, I.; Crawford, A.W.; Michelsen, J.W.; Beckerle, M.C.  
J. Cell Biol. 119, 1573-1587, 1992  
A;Title: Zyxin and cCRP: two interactive LIM domain proteins associated with the  
cytoskeleton.  
A;Reference number: A44358; MUID:93107157; PMID:1469049  
A;Accession: B44358  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-58 <SAD>  
A;Note: sequence extracted from NCBI backbone (NCBIP:121176)  
C;Superfamily: mouse muscle LIM protein MLP; LIM metal-binding repeat homology  
F;9-58/Domain: LIM metal-binding repeat homology (fragment) <LIM>

Query Match 70.0%; Score 28; DB 2; Length 58;  
Best Local Similarity 62.5%; Pred. No. 35;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| |::||:  
Db 14 KAVYFAEE 21

#### RESULT 75

A05035

translation initiation factor IF-1 homolog - common tobacco chloroplast  
(fragment)

N;Alternate names: hypothetical protein 96

C;Species: chloroplast Nicotiana tabacum (common tobacco)

C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 20-Jun-2000

C;Accession: A05035

R;Sugiura, M.

submitted to the EMBL Data Library, August 1986

A;Reference number: A00149

A;Accession: A05035

A;Molecule type: DNA

A;Residues: 1-96 <SUG>

A;Cross-references: EMBL:Z00044; NID:g11807; PID:g1223659

A;Experimental source: cv. Bright Yellow 4

R;Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashida, N.;  
Matsubayashi, T.; Zaita, N.; Chunwongse, J.; Obokata, J.; Yamaguchi-Shinozaki,  
K.; Ohto, C.; Torazawa, K.; Meng, B.Y.; Sugita, M.; Deno, H.; Kamogashira, T.;  
Yamada, K.; Kusuda, J.; Takaiwa, F.; Kato, A.; Tohdoh, N.; Shimada, H.; Sugiura,  
M.

EMBO J. 5, 2043-2049, 1986

A;Title: The complete nucleotide sequence of the tobacco chloroplast genome: its  
gene organization and expression.

A;Reference number: A38013

A;Contents: annotation; gene organization, sites, features

C;Genetics:

A;Genome: chloroplast

C;Superfamily: translation initiation factor IF-1

C;Keywords: chloroplast; protein biosynthesis

Query Match 70.0%; Score 28; DB 2; Length 96;  
Best Local Similarity 71.4%; Pred. No. 58;



Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
 ||:|| |  
 Db 84 KLIFQEQ 90

RESULT 76

B64450

ribosomal protein HS6-type - *Methanococcus jannaschii*

C;Species: *Methanococcus jannaschii*

C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 26-Feb-1998

C;Accession: B64450

R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton, R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.; Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Scott, J.L.; Geoghagen, N.S.M.; Weidman, J.F.; Fuhrmann, J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*.

A;Reference number: A64300; MUID:96337999; PMID:8688087

A;Accession: B64450

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-120 <BUL>

A;Cross-references: GB:U67562; GB:L77117; NID:g1591830; PID:g1591833;

TIGR:MJ1203; PID:g1511204

C;Genetics:

A;Map position: REV1147262-1146900

A;Start codon: TTG

C;Superfamily: rat ribosomal protein L7a

Query Match 70.0%; Score 28; DB 2; Length 120;  
 Best Local Similarity 75.0%; Pred. No. 73;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||| |||  
 Db 47 KLVIIAED 54

RESULT 77

D71025

probable ribosomal protein L7 - *Pyrococcus horikoshii*

C;Species: *Pyrococcus horikoshii*

C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000

C;Accession: D71025

R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine, M.; Baba, S.; Kosugi, H.; Hosoyama, A.; Nagai, Y.; Sakai, M.; Ogura, K.; Otsuka, R.; Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.; Aoki, K.; Yoshizawa, T.; Nakamura, Y.; Robb, F.T.; Horikoshi, K.; Masuchi, Y.; Shizuya, H.; Kikuchi, H.

DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3.

A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: D71025

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-124 <KAW>

A;Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30604.1; PID:g3257921

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:

A;Gene: PH1496

C;Superfamily: rat ribosomal protein L7a

Query Match	70.0%;	Score 28;	DB 2;	Length 124;
Best Local Similarity	75.0%;	Pred. No. 75;		
Matches	6;	Conservative	0;	Mismatches 2; Indels 0; Gaps 0;

Qy	1	KLVFFAED	8
Db	50	KLVIIAED	57

#### RESULT 78

C75109

lsu ribosomal protein l7ae (rpl7ae) PAB0460 - *Pyrococcus abyssi* (strain Orsay)

C;Species: *Pyrococcus abyssi*

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C;Accession: C75109

R;anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure and evolution.

A;Reference number: A75001

A;Accession: C75109

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-128 <KAW>

A;Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49588.1; PID:g5458099

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: rpl7AE; PAB0460

C;Superfamily: rat ribosomal protein L7a

Query Match	70.0%;	Score 28;	DB 2;	Length 128;
Best Local Similarity	75.0%;	Pred. No. 78;		
Matches	6;	Conservative	0;	Mismatches 2; Indels 0; Gaps 0;

Qy	1	KLVFFAED	8
Db	54	KLVIIAED	61

#### RESULT 79

S75397

ribosomal protein HS6 - *Sulfolobus solfataricus*

N;Alternate names: protein c04031

C;Species: *Sulfolobus solfataricus*

C;Date: 09-Oct-1997 #sequence\_revision 24-Oct-1997 #text\_change 20-Jun-2000

C;Accession: S75397

R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.L.; Young, F.; Schenk, M.E.; Gaasterland, T.; Doolittle, W.F.; Ragan, M.A.; Charlebois, R.L.

Mol. Microbiol. 22, 175-191, 1996

A;Title: Organizational characteristics and information content of an archaeal genome: 156 kb of sequence from *Sulfolobus solfataricus* P2.

A;Reference number: S73076; MUID:97055432; PMID:8899719

A;Accession: S75397

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-130 <SEN>

A;Cross-references: EMBL:Y08257; NID:g1707772; PIDN:CAA69560.1; PID:g1707803

A;Experimental source: strain P2

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996

C;Superfamily: rat ribosomal protein L7a

C;Keywords: protein biosynthesis; ribosome

Query Match 70.0%; Score 28; DB 2; Length 130;  
Best Local Similarity 75.0%; Pred. No. 79;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||| |||  
Db 52 KLVIIAED 59

RESULT 80

S67150

hypothetical protein YOR253w - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: hypothetical protein O5315

C;Species: *Saccharomyces cerevisiae*

C;Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 19-Apr-2002

C;Accession: S67150

R;Jauniaux, J.C.; Poirey, R.

submitted to the Protein Sequence Database, July 1996

A;Reference number: S67143

A;Accession: S67150

A;Molecule type: DNA

A;Residues: 1-176 <JAU>

A;Cross-references: EMBL:Z75161; NID:g1420572; PID:e252411; PID:g1420573;

GSPDB:GN00015; MIPS:YOR253w

A;Experimental source: strain S288C

C;Genetics:

A;Gene: MIPS:YOR253w

A;Cross-references: SGD:S0005779

A;Map position: 15R

Query Match 70.0%; Score 28; DB 2; Length 176;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||: ||||  
Db 110 KLLKFAED 117

RESULT 81

D81269

hypothetical protein Cj1712 [imported] - Campylobacter jejuni (strain NCTC 11168)

C;Species: Campylobacter jejuni

C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002

C;Accession: D81269

R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R.M.; Feltwell, T.; Holroyd, S.; Jagels, K.; Karlyshev, A.; Moule, S.; Pallen, M.J.; Penn, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.G.  
Nature 403, 665-668, 2000

A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences.

A;Reference number: A81250; MUID:20150912; PMID:10688204

A;Accession: D81269

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-178 <PAR>

A;Cross-references: GB:AL139079; GB:AL111168; NID:g6968971; PIDN:CAB73698.1; PID:g6969127; GSPDB:GN00120; CJSP:Cj1712

A;Experimental source: serotype O2, strain NCTC 11168

C;Genetics:

A;Gene: Cj1712

Query Match	70.0%;	Score 28;	DB 2;	Length 178;
Best Local Similarity	75.0%;	Pred. No. 1.1e+02;		
Matches	6;	Conservative	1;	Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 || ||||:  
Db 115 KLGFFAEN 122

RESULT 82

T30677

hypothetical protein 75R - Molluscum contagiosum virus 1

N;Alternate names: MC075R

C;Species: Molluscum contagiosum virus 1

C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 11-May-2000

C;Accession: T30677

R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.  
Science 273, 813-816, 1996

A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host response-evasion genes.

A;Reference number: Z20876; MUID:96325459; PMID:8670425

A;Accession: T30677

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-183 <SEN>

A;Cross-references: EMBL:U60315; PIDN:AAC55203.1

C;Genetics:  
A;Note: MC075R

Query Match 70.0%; Score 28; DB 2; Length 183;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
:|||||  
Db 10 MFFAED 15

RESULT 83

S38879

LIM-domain protein CRP1 - Japanese quail

C;Species: Coturnix coturnix japonica (Japanese quail)

C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 15-Jun-2001

C;Accession: S38879

R;Weiskirchen, R.; Beckerle, M.C.; Bister, K.

submitted to the EMBL Data Library, November 1993

A;Description: The CRP family of LIM-domain proteins: identification of two different and independently regulated crp genes in avian species.

A;Reference number: S38879

A;Accession: S38879

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-192 <WEI>

A;Cross-references: EMBL:Z28333; NID:g429159; PID:g1334627

C;Superfamily: mouse muscle LIM protein MLP; LIM metal-binding repeat homology

F;10-61/Domain: LIM metal-binding repeat homology <LIM1>

F;118-169/Domain: LIM metal-binding repeat homology <LIM2>

Query Match 70.0%; Score 28; DB 2; Length 192;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| |:|||:  
Db 15 KAVYFAEE 22

RESULT 84

A49648

cysteine-rich protein - chicken

N;Alternate names: zyxin-binding protein

C;Species: Gallus gallus (chicken)

C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 15-Jun-2001

C;Accession: A49648; S34179

R;Crawford, A.W.; Pino, J.D.; Beckerle, M.C.

J. Cell Biol. 124, 117-127, 1994

A;Title: Biochemical and molecular characterization of the chicken cysteine-rich protein, a developmentally regulated LIM-domain protein that is associated with the actin cytoskeleton.

A;Reference number: A49648; MUID:94124603; PMID:8294495

A;Accession: A49648

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-192 <CRA>  
A;Cross-references: EMBL:X73831; NID:g313286; PIDN:CAA52053.1; PID:g313287  
C;Superfamily: mouse muscle LIM protein MLP; LIM metal-binding repeat homology  
C;Keywords: actin binding; muscle  
F;10-61/Domain: LIM metal-binding repeat homology <LIM1>  
F;118-169/Domain: LIM metal-binding repeat homology <LIM2>

Query Match 70.0%; Score 28; DB 2; Length 192;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | : | | :  
Db 15 KAVYFAEE 22

RESULT 85

S53580

cysteine-rich protein - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 15-Jun-2001

C;Accession: S53580

R;McLaughlin, C.R.; Tao, Q.; Abood, M.E.

Nucleic Acids Res. 22, 5477-5483, 1994

A;Title: Isolation and developmental expression of a rat cDNA encoding a putative cysteine-rich zinc finger protein.

A;Reference number: S53580; MUID:95116342; PMID:7816640

A;Accession: S53580

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-193 <MCL>

A;Cross-references: EMBL:U09567; NID:g563809; PIDN:AAC52157.1; PID:g563810

C;Superfamily: mouse muscle LIM protein MLP; LIM metal-binding repeat homology

F;10-61/Domain: LIM metal-binding repeat homology <LIM1>

F;119-170/Domain: LIM metal-binding repeat homology <LIM2>

Query Match 70.0%; Score 28; DB 2; Length 193;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | : | | :  
Db 15 KTVYFAEE 22

RESULT 86

S12658

cysteine-rich protein - human

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 15-Jun-2001

C;Accession: S12658; A40194

R;Liebhaber, S.A.; Emery, J.G.; Urbanek, M.; Wang, X.; Cooke, N.E.

Nucleic Acids Res. 18, 3871-3879, 1990

A;Title: Characterization of a human cDNA encoding a widely expressed and highly conserved cysteine-rich protein with an unusual zinc-finger motif.

A;Reference number: S12658; MUID:90326508; PMID:2115670

A;Accession: S12658

A;Molecule type: mRNA  
 A;Residues: 1-193 <LIE>  
 A;Cross-references: EMBL:M33146; NID:g181070; PIDN:AAA35720.1; PID:g181071  
 R;Wang, X.; Lee, G.; Liebhaber, S.A.; Cooke, N.E.  
 J. Biol. Chem. 267, 9176-9184, 1992  
 A;Title: Human cysteine-rich protein. A member of the LIM/double-finger family displaying coordinate serum induction with c-myc.  
 A;Reference number: A40194; MUID:92250516; PMID:1374386  
 A;Accession: A40194  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-193 <WAN>  
 A;Cross-references: GB:M76375; GB:M76376; GB:M76377; GB:M76378; NID:g181063; PID:g181064  
 C;Genetics:  
 A;Gene: GDB:CSRPl; CSRP; D1S181E  
 A;Cross-references: GDB:135005; OMIM:123876  
 A;Map position: 1q24-1q32  
 C;Superfamily: mouse muscle LIM protein MLP; LIM metal-binding repeat homology  
 C;Keywords: DNA binding; zinc finger  
 F;10-61/Domain: LIM metal-binding repeat homology <LIM1>  
 F;119-170/Domain: LIM metal-binding repeat homology <LIM2>

Query Match 70.0%; Score 28; DB 2; Length 193;  
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 | | : | | :  
 Db 15 KTVYFAEE 22

#### RESULT 87

T24709  
 hypothetical protein T08G5.1 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T24709  
 R;Smye, R.  
 submitted to the EMBL Data Library, November 1996  
 A;Reference number: Z19927  
 A;Accession: T24709  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-272 <WIL>  
 A;Cross-references: EMBL:Z81589; PIDN:CAB04717.1; GSPDB:GN00023; CESP:T08G5.1  
 A;Experimental source: clone T08G5  
 C;Genetics:  
 A;Gene: CESP:T08G5.1  
 A;Map position: 5  
 A;Introns: 15/1; 38/1; 179/2; 207/3; 241/2

Query Match 70.0%; Score 28; DB 2; Length 272;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8

Db           |::|| |:  
246 KIIFFTEN 253

RESULT 88

T32480

hypothetical protein F52H2.5 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 18-Feb-2000

C;Accession: T32480

R;Chisoe, S.; Hawkins, J.

submitted to the EMBL Data Library, September 1997

A;Description: The sequence of *C. elegans* cosmid F52H2.

A;Reference number: Z21175

A;Accession: T32480

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-283 <CHI>

A;Cross-references: EMBL:AF026214; PIDN:AAB71315.1; GSPDB:GN00028; CESP:F52H2.5

A;Experimental source: strain Bristol N2; clone F52H2

C;Genetics:

A;Gene: CESP:F52H2.5

A;Map position: X

A;Introns: 38/1; 73/2; 146/3; 226/2

Query Match           70.0%;   Score 28;   DB 2;   Length 283;  
Best Local Similarity   62.5%;   Pred. No. 1.7e+02;  
Matches       5;   Conservative       1;   Mismatches       2;   Indels       0;   Gaps       0;

Qy           1 KLVFFAED 8  
             |:| | ||  
Db           80 KIVIFVED 87

RESULT 89

AB1397

hypothetical protein lmo2578 [imported] - *Listeria monocytogenes* (strain EGD-e)

C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C;Accession: AB1397

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AB1397

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-291 <GLA>



A;Cross-references: GB:NC\_003210; PIDN:CAD00656.1; PID:g16412066; GSPDB:GN00177  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmo2578

Query Match 70.0%; Score 28; DB 2; Length 291;  
Best Local Similarity 71.4%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
| :|||  
Db 145 LTYFAED 151

RESULT 90

A36966

probable dTDP-rhamnosyl transferase RfbF - *Shigella flexneri*

C;Species: *Shigella flexneri*

C;Date: 13-Sep-1995 #sequence\_revision 13-Sep-1995 #text\_change 08-Oct-1999

C;Accession: A36966

R;Morona, R.; Mavris, M.; Fallarino, A.; Manning, P.A.

J. Bacteriol. 176, 733-747, 1994

A;Title: Characterization of the rfc region of *Shigella flexneri*.

A;Reference number: A36966; MUID:94131953; PMID:7507920

A;Accession: A36966

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-296 <MOR>

A;Cross-references: GB:X71970; NID:g506557; PIDN:CAA50772.1; PID:g454903

C;Genetics:

A;Gene: rfbF

Query Match 70.0%; Score 28; DB 2; Length 296;  
Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| ::|| :|  
Db 83 KIIFFDQD 90

RESULT 91

AI1160

CheA activity-modulating chemotaxis protein CheV homolog lmo0689 [imported] -  
*Listeria monocytogenes* (strain EGD-e)

C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C;Accession: AI1160

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;  
Bloeker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve,  
E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.;  
Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.;  
Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.;  
Jackson, D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.;  
Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos,

B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.  
A;Title: Comparative genomics of *Listeria* species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AI1160  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-302 <GLA>  
A;Cross-references: GB:NC\_003210; PIDN:CAC98767.1; PID:g16410078; GSPDB:GN00177  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmo0689  
C;Superfamily: chemotaxis cheV protein; response regulator homology

Query Match 70.0%; Score 28; DB 2; Length 302;  
Best Local Similarity 62.5%; Pred. No. 1.9e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| :| |||  
Db 174 KTIFIAED 181

#### RESULT 92

AI1519

CheA activity-modulating chemotaxis protein CheV homolog lin0697 [imported] -  
*Listeria innocua* (strain Clip11262)

C;Species: *Listeria innocua*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C;Accession: AI1519

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AI1519

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-302 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC95929.1; PID:g16413149; GSPDB:GN00178

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: lin0697

C;Superfamily: chemotaxis cheV protein; response regulator homology

Query Match 70.0%; Score 28; DB 2; Length 302;  
Best Local Similarity 62.5%; Pred. No. 1.9e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| :| |||  
Db 174 KTIFIAED 181

RESULT 93

T45723

hypothetical protein F1P2.180 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 15-Sep-2000

C;Accession: T45723

R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.;  
Artiguenave, F.; Saurin, W.; Weissenbach, J.; Mewes, H.W.; Mayer, K.F.X.;  
Lemcke, K.; Schueller, C.; Quetier, F.; Salanoubat, M.

submitted to the Protein Sequence Database, November 1999

A;Reference number: Z23010

A;Accession: T45723

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-332 <CHO>

A;Cross-references: EMBL:AL132955

A;Experimental source: cultivar Columbia; BAC clone F1P2

C;Genetics:

A;Map position: 3

A;Introns: 36/3; 55/3; 76/3; 99/3; 130/3; 165/3; 191/3; 234/3; 270/1

A;Note: F1P2.180

C;Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC1A4.06c

Query Match 70.0%; Score 28; DB 2; Length 332;  
Best Local Similarity 83.3%; Pred. No. 2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
:|||||  
Db 183 MFFAED 188

RESULT 94

T20562

hypothetical protein F07H5.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T20562

R;Steward, C.

submitted to the EMBL Data Library, December 1995

A;Reference number: Z19292

A;Accession: T20562

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-334 <WIL>

A;Cross-references: EMBL:Z68314; PIDN:CAA92663.1; GSPDB:GN00020; CESP:F07H5.2

A;Experimental source: clone F07H5

C;Genetics:

A;Gene: CESP:F07H5.2

A;Map position: 2

A;Introns: 72/2; 146/3; 217/1; 280/3

Query Match 70.0%; Score 28; DB 2; Length 334;  
Best Local Similarity 71.4%; Pred. No. 2.1e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
||:||||:  
Db 196 KLLFFAD 202

RESULT 95

G82596

phage-related protein XF2127 [imported] - *Xylella fastidiosa* (strain 9a5c)

C;Species: *Xylella fastidiosa*

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C;Accession: G82596

R;anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis, Sao Paulo, Brazil.

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: G82596

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-339 <SIM>

A;Cross-references: GB:AE004027; GB:AE003849; NID:g9107249; PIDN:AAF84926.1; GSPDB:GN00128; XFSC:XF2127

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Alves, L.M.C.; Araya, J.E.; Baia, G.S.; Baptista, C.S.; Barros, M.H.; Bonaccorsi, E.D.; Bordin, S.; Bove, J.M.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.; Colauto, N.B.; Colombo, C.; Costa, F.F.; Costa, M.C.R.; Costa-Neto, C.M.; Coutinho, L.L.; Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme, M.; Furlan, L.R.; Garnier, M.; Goldman, G.H.; Goldman, M.H.S.; Gomes, S.L.; Gruber, A.; Ho, P.L.; Hoheisel, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret, F.; Lambais, M.R.; Leite, L.C.C.; Lemos, E.G.M.; Lemos, M.V.F.; Lopes, S.A.; Lopes, C.R.; Machado, J.A.; Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.L.

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; Monteiro-Vitorello, C.B.; Moon, D.H.; Nagai, M.A.; Nascimento, A.L.T.O.; Netto, L.E.S.; Nhani Jr., A.; Nobrega, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Paris, A.; Peixoto, B.R.; Pereira, G.A.G.; Pereira Jr., H.A.; Pesquero, J.B.; Quaggio, R.B.; Roberto, P.G.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, H.E.

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira, J.F.; Silvestri, M.L.Z.; Siqueira, W.J.; de Souza, A.A.; de Souza, A.P.; Terenzi, M.F.; Truffi, D.; Tsai, S.M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zago, M.A.; Zatz, M.

Meidanis, J.; Setubal, J.C.

A;Reference number: A59328

A;Contents: annotation

C;Genetics:  
A;Gene: XF2127

Query Match 70.0%; Score 28; DB 2; Length 339;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| |||| |  
Db 297 KRVFFAAD 304

RESULT 96

AE2425

hypothetical protein alr4957 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C;Accession: AE2425

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.;  
Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.;  
Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.;  
Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing  
Cyanobacterium Anabaena sp. strain PCC 7120.

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AE2425

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-358 <KUR>

A;Cross-references: GB:BA000019; PIDN:BA076656.1; PID:g17134095; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr4957

C;Superfamily: conserved hypothetical protein yacL

Query Match 70.0%; Score 28; DB 2; Length 358;  
Best Local Similarity 75.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||| ||:|  
Db 245 KLVRF AQD 252

RESULT 97

A96535

unknown protein, 11341-13243 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 28-Jul-2003

C;Accession: A96535

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;  
Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.;  
Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.;  
Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.;

Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda, M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: A96535  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-420 <STO>  
 A;Cross-references: GB:AE005173; NID:g10092169; PIDN:AAG12589.1; GSPDB:GN00141  
 C;Genetics:  
 A;Gene: F10F5.1  
 A;Map position: 1  
 C;Superfamily: 5-methylthioribose kinase

Query Match 70.0%; Score 28; DB 2; Length 420;  
 Best Local Similarity 71.4%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
 | |||:  
 Db 277 LAFFAQD 283

# RESULT 98

G81271  
 biotin carboxylase (EC 6.3.4.14) Cj1290c [imported] - Campylobacter jejuni (strain NCTC 11168)  
 C;Species: Campylobacter jejuni  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C;Accession: G81271  
 R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R.M.; Feltwell, T.; Holroyd, S.; Jagels, K.; Karlyshev, A.; Moule, S.; Pallen, M.J.; Penn, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.G.  
 Nature 403, 665-668, 2000  
 A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences.  
 A;Reference number: A81250; MUID:20150912; PMID:10688204  
 A;Accession: G81271  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-443 <PAR>  
 A;Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73717.1; PID:g6968724; GSPDB:GN00120; CJSP:Cj1290c  
 A;Experimental source: serotype O2, strain NCTC 11168  
 C;Genetics:

A;Gene: accC; Cj1290c  
C;Superfamily: biotin carboxylase; biotin carboxylase homology  
C;Keywords: ligase

Query Match 70.0%; Score 28; DB 2; Length 443;  
Best Local Similarity 75.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||| :|||  
Db 387 KLVVWAED 394

RESULT 99

T02324

hypothetical protein At2g34360 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein F13P17.20

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Feb-2001

C;Accession: T02324; F84755

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.; Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter, J.C.

submitted to the EMBL Data Library, July 1998

A;Description: Arabidopsis thaliana chromosome II BAC F13P17 genomic sequence.

A;Reference number: Z14657

A;Accession: T02324

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-466 <ROU>

A;Cross-references: EMBL:AC004481; NID:g3337347; PID:g3337367

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: F84755

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-466 <STO>

A;Cross-references: GB:AE002093; NID:g3337367; PIDN:AAC27412.1; GSPDB:GN00139

C;Genetics:

A;Gene: F13P17.20; At2g34360

A;Map position: 2

A;Introns: 82/3; 264/2; 292/1; 306/2; 386/1; 425/3

C;Superfamily: conserved hypothetical protein HI1612

Query Match 70.0%; Score 28; DB 2; Length 466;  
Best Local Similarity 71.4%; Pred. No. 2.9e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
 |||| :|  
Db 134 LVFFGQD 140

RESULT 100

AI3442

glycolate oxidase chain glcD [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002

C;Accession: AI3442

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.;

Ivanova, N.; Anderson, I.; Bhattacharyya, A.; Lykidis, A.; Reznik, G.;

Jablonski, L.; Larsen, N.; D'Souza, M.; Bernal, A.; Mazur, M.; Goltsman, E.;

Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesson, J.J.; Haselkorn,

R.; Kyrpides, N.; Overbeek, R.

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis.

A;Reference number: AD3252; PMID:11756688

A;Accession: AI3442

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-470 <KUR>

A;Cross-references: GB:AE008917; PIDN:AAL52708.1; PID:g17983536; GSPDB:GN00190

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEI1527

A;Map position: I

C;Superfamily: glycolate oxidase chain glcD

Query Match 70.0%; Score 28; DB 2; Length 470;

Best Local Similarity 62.5%; Pred. No. 2.9e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 :|||| :|  
Db 438 ELVFFKQD 445

Search completed: February 28, 2004, 08:56:36

Job time : 94 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 28, 2004, 07:06:39 ; Search time 21.5 Seconds  
(without alignments)  
19.375 Million cell updates/sec

Title: US-09-668-314C-73  
Perfect score: 40  
Sequence: 1 KLVFFAED 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	100.0	57	1	A4_URSMA	Q29149 ursus marit
2	40	100.0	58	1	A4_CANFA	Q28280 canis famil
3	40	100.0	58	1	A4_RABIT	Q28748 oryctolagus
4	40	100.0	58	1	A4_SHEEP	Q28757 ovis aries
5	40	100.0	59	1	A4_BOVIN	Q28053 bos taurus
6	40	100.0	751	1	A4_SAIISC	Q95241 s amyloid b
7	40	100.0	770	1	A4_CAVPO	Q60495 c amyloid b
8	40	100.0	770	1	A4_HUMAN	P05067 h amyloid b
9	40	100.0	770	1	A4_MACFA	P53601 m amyloid b
10	40	100.0	770	1	A4_MOUSE	P12023 m amyloid b
11	40	100.0	770	1	A4_PIG	P79307 s amyloid b
12	40	100.0	770	1	A4_RAT	P08592 r amyloid b
13	40	100.0	780	1	A4_TETFL	O73683 tetraodon f
14	37	92.5	737	1	A4_FUGRU	O93279 fugu rubrip
15	34	85.0	321	1	Y189_RICPR	Q9zdx5 rickettsia
16	31	77.5	119	1	RL7A_ARCFU	O29494 archaeoglob
17	31	77.5	281	1	UPK_CORST	Q9fb58 corynebacte

18	31	77.5	641	1	LICR_BACSU	P46321	bacillus su
19	31	77.5	2196	1	MOR2_SCHPO	Q9hdv6	schizosacch
20	30	75.0	214	1	UT11_ORYSA	Q8slz1	oryza sativ
21	30	75.0	341	1	Y665_METJA	Q58079	methanococc
22	30	75.0	370	1	GP85_HUMAN	Q9npd1	homo sapien
23	30	75.0	371	1	GP85_BRARE	Q9i919	brachydanio
24	30	75.0	533	1	SYK_METMP	O30522	methanococc
25	30	75.0	1353	1	CYA9_HUMAN	O60503	homo sapien
26	30	75.0	1353	1	CYA9_MOUSE	P51830	mus musculu
27	29	72.5	119	1	RL7A_HALMA	P12743	haloarcula
28	29	72.5	189	1	CME2_BACSU	P32393	bacillus su
29	29	72.5	258	1	RSFA_BACSU	P39650	bacillus su
30	29	72.5	278	1	ERA_BUCAP	Q8k9r2	buchnera ap
31	29	72.5	286	1	DMA_HAEIN	P44431	haemophilus
32	29	72.5	357	1	HST2_YEAST	P53686	saccharomyc
33	29	72.5	469	1	RECJ_METJA	Q58387	methanococc
34	29	72.5	481	1	RNF9_PANTR	Q7yr32	pan troglod
35	29	72.5	482	1	RNF9_HUMAN	Q9udy6	homo sapien
36	29	72.5	634	1	ELM1_ASPFU	P46074	aspergillus
37	29	72.5	634	1	ELM2_ASPFU	P46075	aspergillus
38	29	72.5	809	1	KCB2_CANFA	Q95167	canis famil
39	29	72.5	857	1	KCB1_MOUSE	Q03717	mus musculu
40	29	72.5	857	1	KCB1_RAT	P15387	rattus norv
41	29	72.5	858	1	KCB1_HUMAN	Q14721	homo sapien
42	29	72.5	858	1	KCB1_PIG	O18868	sus scrofa
43	29	72.5	858	1	KCB1_RABIT	Q9mz19	oryctolagus
44	29	72.5	907	1	KCB2_RAT	Q63099	rattus norv
45	29	72.5	911	1	KCB2_HUMAN	Q92953	homo sapien
46	29	72.5	911	1	KCB2_RABIT	Q95111	oryctolagus
47	29	72.5	3080	1	POLG_ZYMVC	P18479	z genome po
48	29	72.5	3083	1	POLG_ZYMVS	O36979	z genome po
49	29	72.5	4466	1	DYHC_ANTCR	P39057	anthocidari
50	29	72.5	4466	1	DYHC_TRIGR	P23098	tripneustes
51	28	70.0	96	1	IF1C_TOBAC	P12136	nicotiana t
52	28	70.0	117	1	RL7A_METJA	P54066	methanococc
53	28	70.0	120	1	RL7A_METAC	Q8tql9	methanosarc
54	28	70.0	120	1	RL7A_METMA	Q8pu74	methanosarc
55	28	70.0	123	1	RL7A_PYRHO	O59165	pyrococcus
56	28	70.0	124	1	RL7A_PYRFU	Q8ul60	pyrococcus
57	28	70.0	127	1	RL7A_SULSO	P55858	sulfolobus
58	28	70.0	151	1	RL7A_PYRAE	Q8zta5	pyrobaculum
59	28	70.0	191	1	CYSR_CHICK	P32965	gallus gall
60	28	70.0	192	1	CYSR_HUMAN	P21291	homo sapien
61	28	70.0	192	1	CYSR_RAT	P47875	rattus norv
62	28	70.0	296	1	RFBF_SHIFL	P37782	shigella fl
63	28	70.0	417	1	O65A_DROME	P82982	drosophila
64	28	70.0	444	1	BCHN_CHLAU	Q9f6x6	chloroflexu
65	28	70.0	495	1	CRTD_RHOSH	Q01671	rhodobacter
66	28	70.0	540	1	CH60_CLOTM	P48212	clostridium
67	28	70.0	551	1	YABN_ECOLI	P33595	escherichia
68	28	70.0	575	1	APB3_HUMAN	O96018	homo sapien
69	28	70.0	623	1	HCYE_EURCA	P02242	eurypelma c
70	28	70.0	628	1	HCYF_EURCA	Q9nfl5	eurypelma c
71	28	70.0	628	1	HCYG_EURCA	Q9nfl4	eurypelma c
72	28	70.0	630	1	HCYA_EURCA	P14750	eurypelma c
73	28	70.0	685	1	MDL1_CANAL	P97998	candida alb
74	28	70.0	728	1	VIV_ORYSA	P37398	oryza sativ

75	28	70.0	754	1	CHLD_PEA	022437	pisum sativ
76	28	70.0	758	1	CHLD_TOBAC	024133	nicotiana t
77	28	70.0	759	1	CHLD_ARATH	Q9sje1	arabidopsis
78	28	70.0	3898	1	POLG_HCVB	P21530	hog cholera
79	28	70.0	4486	1	DYH9_HUMAN	Q9nyc9	homo sapien
80	28	70.0	4523	1	DYHB_HUMAN	Q96dt5	homo sapien
81	27	67.5	109	1	VNS2_CVBM	P15774	bovine coro
82	27	67.5	109	1	VNS2_CVHOC	Q04853	human coron
83	27	67.5	121	1	RL7A_THEAC	Q9hj56	thermoplasm
84	27	67.5	121	1	RL7A_THEVO	Q97bk8	thermoplasm
85	27	67.5	127	1	RL7A_AERPE	Q9yax7	aeropyrum p
86	27	67.5	204	1	YUFG_BACSU	O05249	bacillus su
87	27	67.5	258	1	TATC_ECOLI	P27857	escherichia
88	27	67.5	271	1	YDJO_ECOLI	P76210	escherichia
89	27	67.5	282	1	UPK_CHLTE	Q8kfj7	chlorobium
90	27	67.5	293	1	UPK_CORGL	Q8nqc3	corynebacte
91	27	67.5	293	1	XT31_ARATH	P93046	arabidopsis
92	27	67.5	293	1	XTH5_ARATH	Q9xiw1	arabidopsis
93	27	67.5	293	1	Y347_HELPJ	Q9zma1	helicobacte
94	27	67.5	305	1	GBB_SCHPO	Q10282	schizosacch
95	27	67.5	338	1	AHA1_HUMAN	O95433	homo sapien
96	27	67.5	338	1	AHA1_MOUSE	Q8bk64	mus musculu
97	27	67.5	387	1	YGH5_YEAST	P53159	saccharomyc
98	27	67.5	387	1	YM14_PARTE	P15615	paramecium
99	27	67.5	394	1	AV71_ACAVI	Q17107	acanthochei
100	27	67.5	432	1	OV71_ONCVO	P31732	onchocerca
101	27	67.5	437	1	DNAA_MYCGE	P35888	mycoplasma
102	27	67.5	438	1	ENGA_CLOPE	Q8xjk1	clostridium
103	27	67.5	465	1	RBL_NEPAL	P28434	nepenthes a
104	27	67.5	468	1	LIP1_CANAL	O94091	candida alb
105	27	67.5	492	1	YD48_SCHPO	Q10301	schizosacch
106	27	67.5	500	1	YJGR_ECOLI	P39342	escherichia
107	27	67.5	580	1	SRK1_SCHPO	O94547	schizosacch
108	27	67.5	610	1	RCK2_YEAST	P38623	saccharomyc
109	27	67.5	631	1	PRKA_BACSU	P39134	bacillus su
110	27	67.5	772	1	LP1G_DROME	P11997	drosophila
111	27	67.5	789	1	LP1A_DROME	P11995	drosophila
112	27	67.5	789	1	LP1B_DROME	P11996	drosophila
113	27	67.5	930	1	SYI_STRPN	Q9zhh3	streptococc
114	27	67.5	956	1	MTN2_HUMAN	O00339	homo sapien
115	27	67.5	956	1	MTN2_MOUSE	O08746	mus musculu
116	27	67.5	1713	1	LMA3_HUMAN	Q16787	homo sapien
117	27	67.5	2164	1	CCAA_MOUSE	P97445	mus musculu
118	27	67.5	2184	1	RRPL_CDVO	P24658	canine dist
119	27	67.5	3068	1	POLG_PEMVC	Q01500	p genome po
120	27	67.5	3137	1	CA36_CHICK	P15989	gallus gall
121	27	67.5	3333	1	LMA3_MOUSE	Q61789	mus musculu
122	27	67.5	5327	1	MACF_MOUSE	Q9qxx0	mus musculu
123	27	67.5	5430	1	MACF_HUMAN	Q9upn3	homo sapien
124	27	67.5	5938	1	MAC4_HUMAN	Q96pk2	homo sapien
125	27	67.5	6629	1	R1AB_IBVB	P27920	a replicase
126	27	67.5	6629	1	R1AB_IBVBC	Q91qt2	a replicase
127	26	65.0	89	1	PE23_SHEEP	Q28550	ovis aries
128	26	65.0	103	1	YLXQ_ENTFC	P55768	enterococcu
129	26	65.0	109	1	Y819_METJA	Q58229	methanococc
130	26	65.0	115	1	YACC_ECOLI	P23838	escherichia
131	26	65.0	120	1	RL7A_HALN1	Q9hgh8	halobacteri

132	26	65.0	139	1	FABZ_STRP3	Q8k631	streptococc
133	26	65.0	139	1	FABZ_STRPY	P58175	streptococc
134	26	65.0	140	1	FABZ_STRMU	Q8dsn8	streptococc
135	26	65.0	140	1	FABZ_STRPN	P59201	streptococc
136	26	65.0	140	1	FABZ_STRR6	P59202	streptococc
137	26	65.0	144	1	FAZ2_LACLA	Q9chf4	lactococcus
138	26	65.0	186	1	GL19_ORYSA	P29835	oryza sativ
139	26	65.0	187	1	Y163_METJA	Q57627	methanococc
140	26	65.0	196	1	YB05_THEAC	Q9hj67	thermoplasm
141	26	65.0	202	1	B3G1_MOUSE	Q9cw73	m galactosy
142	26	65.0	208	1	CRP2_HUMAN	P52943	homo sapien
143	26	65.0	208	1	CRP2_RAT	P36201	rattus norv
144	26	65.0	216	1	LIPB_WIGBR	Q8d326	wiggleswort
145	26	65.0	225	1	FLIH_BUCBP	Q89az8	buchnera ap
146	26	65.0	225	1	YEA1_YEAST	P40006	saccharomyc
147	26	65.0	240	1	Y833_TREPA	O83805	treponema p
148	26	65.0	251	1	Y126_PYRAB	Q9v2e8	pyrococcus
149	26	65.0	255	1	Y439_PYRHO	O58186	pyrococcus
150	26	65.0	257	1	FCEA_HUMAN	P12319	homo sapien
151	26	65.0	267	1	XT11_ARATH	Q9smp1	arabidopsis
152	26	65.0	273	1	LB20_ARATH	Q9srv3	arabidopsis
153	26	65.0	289	1	SGCD_MESAU	P97281	mesocricetu
154	26	65.0	289	1	SGCD_MOUSE	P82347	mus musculu
155	26	65.0	299	1	HIS1_BUCAI	P57200	buchnera ap
156	26	65.0	299	1	HIS1_BUCAP	Q9zhe7	buchnera ap
157	26	65.0	299	1	HIS1_BUCBP	P59453	buchnera ap
158	26	65.0	299	1	HIS1_BUCDN	Q9rq89	buchnera ap
159	26	65.0	299	1	VG12_HSVI1	Q00165	ictalurid h
160	26	65.0	300	1	PRMA_LEPIN	Q8f6b7	leptospira
161	26	65.0	305	1	RPO4_VACCC	P21087	vaccinia vi
162	26	65.0	305	1	RPO4_VACCV	P24757	vaccinia vi
163	26	65.0	305	1	RPO4_VARV	P33812	variola vir
164	26	65.0	306	1	PPT1_MOUSE	O88531	mus musculu
165	26	65.0	306	1	PPT1_RAT	P45479	rattus norv
166	26	65.0	314	1	CRYM_HUMAN	Q14894	homo sapien
167	26	65.0	323	1	B3G2_HUMAN	Q9npz5	homo sapien
168	26	65.0	324	1	B3G2_MOUSE	P59270	mus musculu
169	26	65.0	324	1	B3G2_RAT	Q9z137	rattus norv
170	26	65.0	327	1	RRPP_PIRYV	Q01769	piry virus.
171	26	65.0	334	1	B3G1_HUMAN	Q9p2w7	homo sapien
172	26	65.0	334	1	B3G1_RAT	O35789	r galactosy
173	26	65.0	335	1	B3G3_CRIGR	Q9wu47	cricetulus
174	26	65.0	335	1	B3G3_HUMAN	O94766	homo sapien
175	26	65.0	335	1	B3G3_MOUSE	P58158	mus musculu
176	26	65.0	336	1	USG_AZOVI	P96199	azotobacter
177	26	65.0	340	1	CELF_VZVD	P09261	varicella-z
178	26	65.0	352	1	HIS8_CLOAB	Q97es6	clostridium
179	26	65.0	361	1	Y439_METJA	Q57881	methanococc
180	26	65.0	365	1	Y075_METJA	Q60388	methanococc
181	26	65.0	371	1	RECF_BACHD	Q9rc99	bacillus ha
182	26	65.0	379	1	CYB_MYONA	Q956z9	myotis natt
183	26	65.0	379	1	CYB_PROTA	Q35677	prothecus
184	26	65.0	383	1	O94B_DROME	Q9vcs8	drosophila
185	26	65.0	388	1	ALGJ_AZOVI	O52197	azotobacter
186	26	65.0	400	1	YD01_METJA	Q58697	methanococc
187	26	65.0	403	1	Y271_AQUAE	O66628	aquifex aeo
188	26	65.0	406	1	ARGJ_NEIGO	P38434	n arginine

189	26	65.0	406	1	ARGJ_NEIMA	Q9jrj2	n arginine
190	26	65.0	417	1	PGK_PENCI	P33161	penicillium
191	26	65.0	419	1	HAT1_HUMAN	O14929	homo sapien
192	26	65.0	438	1	CLN3_CANFA	Q29611	canis famil
193	26	65.0	438	1	CLN3_MOUSE	Q61124	mus musculu
194	26	65.0	446	1	AP50_SCHPO	Q09718	schizosacch
195	26	65.0	468	1	SYFA_SULSO	P95961	sulfolobus
196	26	65.0	469	1	MM13_XENLA	Q10835	xenopus lae
197	26	65.0	471	1	SYW_HUMAN	P23381	homo sapien
198	26	65.0	475	1	MM03_RAT	P03957	rattus norv
199	26	65.0	475	1	SYW_BOVIN	P17248	bos taurus
200	26	65.0	475	1	SYW_RABIT	P23612	oryctolagus
201	26	65.0	477	1	MM03_HORSE	Q28397	equus cabal
202	26	65.0	477	1	MM03_HUMAN	P08254	homo sapien
203	26	65.0	477	1	MM03_MOUSE	P28862	mus musculu
204	26	65.0	478	1	MM03_RABIT	P28863	oryctolagus
205	26	65.0	481	1	SYW_MOUSE	P32921	mus musculu
206	26	65.0	481	1	YRDD_CAEEL	Q18953	caenorhabdi
207	26	65.0	490	1	CIKL_DROME	P17971	drosophila
208	26	65.0	493	1	SYE_CORGL	Q8nqx9	corynebacte
209	26	65.0	493	1	SYK_ARCFU	O29052	archaeoglob
210	26	65.0	511	1	VGLG_VSVIG	P04883	vesicular s
211	26	65.0	511	1	VGLG_VSVO	P04884	vesicular s
212	26	65.0	511	1	VGLG_VSVSJ	P03522	vesicular s
213	26	65.0	512	1	CADC_ECOLI	P23890	escherichia
214	26	65.0	513	1	NDI1_YEAST	P32340	saccharomyc
215	26	65.0	518	1	YL92_YEAST	P18634	saccharomyc
216	26	65.0	574	1	HMD3_SOLTU	Q41438	solanum tub
217	26	65.0	575	1	TREZ_ARTRM	Q9ajn6	arthrobacte
218	26	65.0	578	1	MM17_MOUSE	Q9r0s3	mus musculu
219	26	65.0	580	1	MM14_PIG	Q9xt90	sus scrofa
220	26	65.0	582	1	MM14_HUMAN	P50281	homo sapien
221	26	65.0	582	1	MM14_RABIT	Q95220	oryctolagus
222	26	65.0	595	1	HMD2_SOLTU	Q41437	solanum tub
223	26	65.0	595	1	SYD_BACHD	Q9kdgl	bacillus ha
224	26	65.0	596	1	HMD1_SOLTU	P48020	solanum tub
225	26	65.0	602	1	HMD2_LYCES	P48022	lycopersico
226	26	65.0	602	1	SYD_RICCN	Q92j82	rickettsia
227	26	65.0	604	1	HMD2_CAPAN	Q9xel8	capsicum an
228	26	65.0	604	1	HMDH_NICSY	Q01559	nicotiana s
229	26	65.0	605	1	NRDD_BPT4	P07071	bacterioph
230	26	65.0	606	1	MM17_HUMAN	Q9ulz9	homo sapien
231	26	65.0	626	1	HCY6_ANDAU	P80476	androctonus
232	26	65.0	703	1	PURL_SULTO	Q970v6	sulfolobus
233	26	65.0	704	1	SSP2_BOMMO	P20613	bombyx mori
234	26	65.0	747	1	POT1_RAT	Q99pr0	rattus norv
235	26	65.0	749	1	JSB2_TRINI	Q06343	trichoplusi
236	26	65.0	768	1	ITB8_RABIT	P26013	oryctolagus
237	26	65.0	769	1	ITB8_HUMAN	P26012	homo sapien
238	26	65.0	774	1	STF_LAMBD	P03764	bacterioph
239	26	65.0	790	1	SUV6_ARATH	Q8vz17	arabidopsis
240	26	65.0	862	1	LOX1_HORVU	P29114	hordeum vul
241	26	65.0	865	1	LAC9_KLULA	P08657	kluyveromyc
242	26	65.0	901	1	R8L3_ARATH	Q9fjb5	arabidopsis
243	26	65.0	904	1	PPR1_YEAST	P07272	saccharomyc
244	26	65.0	908	1	R8L4_ARATH	Q9fjk8	arabidopsis
245	26	65.0	908	1	RPP8_ARATH	Q8w4j9	arabidopsis

246	26	65.0	910	1	RP8H_ARATH	P59584	arabidopsis
247	26	65.0	919	1	HEX_ADE12	P19900	human adeno
248	26	65.0	969	1	DPOM_NEUIN	P33538	neurospora
249	26	65.0	1016	1	POLG_WMV2U	P18478	w genome po
250	26	65.0	1020	1	YRD3_CAEEL	Q09573	caenorhabdi
251	26	65.0	1033	1	NA11_HUMAN	P59045	homo sapien
252	26	65.0	1045	1	HMD2_YEAST	P12684	saccharomyc
253	26	65.0	1075	1	Y124_METJA	Q57588	methanococc
254	26	65.0	1088	1	SEC8_SCHPO	O74562	schizosacch
255	26	65.0	1144	1	NOS2_MOUSE	P29477	mus musculu
256	26	65.0	1174	1	EX5B_BUCAI	P57529	buchnera ap
257	26	65.0	1217	1	S3B3_HUMAN	Q15393	homo sapien
258	26	65.0	1251	1	YMI1_YEAST	Q04545	saccharomyc
259	26	65.0	1318	1	YDH4_SCHPO	Q92349	schizosacch
260	26	65.0	1325	1	Y309_MYCPN	P75334	mycoplasma
261	26	65.0	1334	1	CYA9_CHICK	Q9dgg6	gallus gall
262	26	65.0	1365	1	SUZ2_DROME	P25172	drosophila
263	26	65.0	1790	1	SEPA_EMENI	P78621	emericella
264	26	65.0	2199	1	DPOE_SCHPO	P87154	schizosacch
265	26	65.0	2366	1	TOXB_CLODI	P18177	clostridium
266	26	65.0	2670	1	IP3T_RAT	Q63269	rattus norv
267	26	65.0	2671	1	IP3T_HUMAN	Q14573	homo sapien
268	26	65.0	3066	1	POLG_BCMVN	Q65399	b genome po
269	26	65.0	3066	1	POLG_SBMVG	Q90069	s genome po
270	26	65.0	3066	1	POLG_SBMVN	P21231	s genome po
271	26	65.0	3083	1	POLG_ZYMVR	Q89330	z genome po
272	25.5	63.7	148	1	LIMA_RHOER	Q9zag3	rhodococcus
273	25	62.5	81	1	MOAD_ECOLI	P30748	escherichia
274	25	62.5	111	1	YG3U_YEAST	P53291	saccharomyc
275	25	62.5	117	1	YCX2_CHLRE	P05722	chlamydomon
276	25	62.5	120	1	WN1A_ALOVU	P28099	alopias vul
277	25	62.5	128	1	Y116_BUCBP	Q89aw2	buchnera ap
278	25	62.5	131	1	Y356_VIBCH	Q9kv02	vibrio chol
279	25	62.5	140	1	FABZ_BACHD	Q9k6j4	bacillus ha
280	25	62.5	153	1	VJ01_VACCC	P21032	vaccinia vi
281	25	62.5	153	1	VJ01_VACCV	P07616	vaccinia vi
282	25	62.5	159	1	ISPF_ECOL6	Q8fej6	escherichia
283	25	62.5	159	1	ISPF_ECOLI	P36663	escherichia
284	25	62.5	159	1	ISPF_SALTI	Q8z472	salmonella
285	25	62.5	159	1	ISPF_SALTY	Q8zmf7	salmonella
286	25	62.5	159	1	VJ01_VARV	P33004	variola vir
287	25	62.5	170	1	APT_FUSNN	Q8rdm9	fusobacteri
288	25	62.5	176	1	RPB7_ARATH	P38421	arabidopsis
289	25	62.5	176	1	RPB7_SOYBN	P46279	glycine max
290	25	62.5	179	1	YD85_METJA	Q58780	methanococc
291	25	62.5	181	1	CS22_BACSU	P94497	bacillus su
292	25	62.5	190	1	RECR_CAMJE	Q9pn35	campylobact
293	25	62.5	194	1	YD16_HAEIN	P44159	haemophilus
294	25	62.5	196	1	BCB1_ARATH	Q07488	arabidopsis
295	25	62.5	196	1	CLPP_PINTH	P41609	pinus thunb
296	25	62.5	205	1	CLPP_PINCO	P36387	pinus conto
297	25	62.5	213	1	YOUB_CAEEL	P34658	caenorhabdi
298	25	62.5	224	1	Y691_CHLTR	O84697	chlamydia t
299	25	62.5	228	1	UT11_ARATH	Q9m223	arabidopsis
300	25	62.5	230	1	PURQ_METJA	Q59042	methanococc
301	25	62.5	235	1	YWFC_BACSU	P39639	bacillus su
302	25	62.5	236	1	Y264_METEX	Q8gek8	methylobact

303	25	62.5	245	1	Y191_LISMO	Q8yae0	listeria mo
304	25	62.5	245	1	Y230_LISIN	Q92f76	listeria in
305	25	62.5	253	1	Y128_PYRHO	O57868	pyrococcus
306	25	62.5	258	1	6PGL_CHLPN	Q9z8u5	chlamydia p
307	25	62.5	260	1	ARGB_BACHD	Q9k8v4	bacillus ha
308	25	62.5	270	1	DHMA_FLAS1	P22441	flavobacter
309	25	62.5	274	1	KSGA_BUCAP	Q8ka00	buchnera ap
310	25	62.5	274	1	UPK1_BACHD	Q9kfl5	bacillus ha
311	25	62.5	276	1	DAPF_VIBPA	Q87kj4	vibrio para
312	25	62.5	278	1	UPK1_STRCO	Q9fc36	streptomyce
313	25	62.5	278	1	UPK2_BACHD	Q9kcp8	bacillus ha
314	25	62.5	280	1	PARB_CHLMU	Q9pln9	chlamydia m
315	25	62.5	281	1	HEMK_AQUAE	O66506	aquifex aeo
316	25	62.5	297	1	FTR_ARCFU	O28076	archaeoglob
317	25	62.5	297	1	FTR_METBA	P55301	methanosarc
318	25	62.5	300	1	ER25_SCHPO	Q9uuh4	schizosacch
319	25	62.5	301	1	YWBI_BACSU	P39592	bacillus su
320	25	62.5	306	1	B3GI_DROME	O97422	d galactosy
321	25	62.5	309	1	Y672_PYRHO	O58405	pyrococcus
322	25	62.5	316	1	DUS_AQUAE	O67533	aquifex aeo
323	25	62.5	321	1	DDL_RICCN	Q92it7	rickettsia
324	25	62.5	321	1	MER_METTH	O27784	methanobact
325	25	62.5	321	1	MER_METTM	Q50744	methanobact
326	25	62.5	332	1	C1B3_CAVPO	Q9qzz0	cavia porce
327	25	62.5	334	1	Y293_BUCAI	P57380	buchnera ap
328	25	62.5	336	1	G3P_ASPNG	Q12552	aspergillus
329	25	62.5	338	1	G3P_ASPOR	Q9hgy7	aspergillus
330	25	62.5	341	1	ACOB_BACSU	O34591	bacillus su
331	25	62.5	343	1	XT30_ARATH	Q38908	arabidopsis
332	25	62.5	350	1	LSS1_MOUSE	P27545	mus musculu
333	25	62.5	357	1	XT29_ARATH	Q817h3	arabidopsis
334	25	62.5	359	1	AG2R_CHICK	P79785	gallus gall
335	25	62.5	359	1	AG2R_MELGA	P33396	meleagris g
336	25	62.5	359	1	MLTA_BUCAI	P57531	buchnera ap
337	25	62.5	361	1	RL4A_YEAST	P10664	saccharomyc
338	25	62.5	361	1	RL4B_YEAST	P49626	saccharomyc
339	25	62.5	367	1	Y584_BUCAI	P57644	buchnera ap
340	25	62.5	369	1	METB_HAEIN	P44502	haemophilus
341	25	62.5	376	1	ACT_ENTHI	P11426	entamoeba h
342	25	62.5	378	1	O33A_DROME	P81914	drosophila
343	25	62.5	379	1	YISP_BACCS	P29156	bacillus cl
344	25	62.5	382	1	P2CA_BOVIN	O62829	bos taurus
345	25	62.5	384	1	ARGD_LISIN	Q92bc0	listeria in
346	25	62.5	385	1	SYW_PYRAB	Q9uy11	pyrococcus
347	25	62.5	386	1	SYW_PYRHO	O59584	pyrococcus
348	25	62.5	401	1	ASSY_AQUAE	O67213	aquifex aeo
349	25	62.5	417	1	YAG1_YEAST	P39713	saccharomyc
350	25	62.5	418	1	NAGA_CAEEL	P34480	caenorhabdi
351	25	62.5	428	1	SYH_BUCAP	Q8k9p3	buchnera ap
352	25	62.5	429	1	GLYA_METJA	Q58992	methanococc
353	25	62.5	434	1	MALE_PYRFU	P58300	pyrococcus
354	25	62.5	443	1	Y314_MYCGE	Q49415	mycoplasma
355	25	62.5	448	1	SDHD_BACSU	P54555	bacillus su
356	25	62.5	453	1	BIOA_AQUAE	O66557	aquifex aeo
357	25	62.5	453	1	MALE_PYRAB	Q9v297	pyrococcus
358	25	62.5	457	1	TRME_LISIN	Q926u7	listeria in
359	25	62.5	457	1	TRME_LISMO	Q8y3m4	listeria mo

360	25	62.5	459	1	RBL_CALSH	P48687	calyptosph
361	25	62.5	459	1	RBL_CHRHI	P48692	chrysochrom
362	25	62.5	464	1	BIOA_METJA	Q58696	methanococc
363	25	62.5	469	1	DLDH_CHLVI	O50311	chlorobium
364	25	62.5	472	1	C901_ARATH	Q42569	arabidopsis
365	25	62.5	473	1	CYSG_BUCAI	P57500	buchnera ap
366	25	62.5	473	1	SYE_WIGBR	Q8d375	wiggleswort
367	25	62.5	474	1	MEC3_YEAST	Q02574	saccharomyc
368	25	62.5	475	1	RBL_ASTLO	P26490	astasia lon
369	25	62.5	480	1	Y066_METJA	Q60377	methanococc
370	25	62.5	482	1	VSAC_TRYBB	P06016	trypanosoma
371	25	62.5	485	1	PUR1_VIGAC	P52419	vigna aconi
372	25	62.5	488	1	RBL_GUIITH	P14957	guillardia
373	25	62.5	488	1	RBL_PLECA	Q08051	pleurochrys
374	25	62.5	489	1	RNF9_MOUSE	Q9wuh5	mus musculu
375	25	62.5	493	1	SYE_COREF	Q8fpu9	corynebacte
376	25	62.5	496	1	AGP2_PIG	Q9bdy7	sus scrofa
377	25	62.5	497	1	MEK1_YEAST	P24719	saccharomyc
378	25	62.5	497	1	RELB_XENLA	P51510	xenopus lae
379	25	62.5	499	1	CIK2_XENLA	P22739	xenopus lae
380	25	62.5	501	1	SCRB_PEDPE	P43471	pediococcus
381	25	62.5	503	1	VSA1_TRYBB	P06015	trypanosoma
382	25	62.5	504	1	SIK1_YEAST	Q12460	saccharomyc
383	25	62.5	509	1	HUTH_PSEPU	P21310	pseudomonas
384	25	62.5	510	1	Y382_RICPR	Q9zde9	rickettsia
385	25	62.5	512	1	FEN2_YEAST	P25621	saccharomyc
386	25	62.5	515	1	SIR2_CANAL	O59923	candida alb
387	25	62.5	518	1	TYD1_PAPSO	P54768	papaver som
388	25	62.5	523	1	TYD5_PAPSO	P54771	papaver som
389	25	62.5	527	1	SYK_RICCN	Q92ib5	rickettsia
390	25	62.5	528	1	SYK_RICPR	Q9zdf8	rickettsia
391	25	62.5	530	1	SYK_METJA	Q57959	methanococc
392	25	62.5	536	1	MVIN_CHLMU	Q9pjb9	chlamydia m
393	25	62.5	536	1	MVIN_CHLTR	Q46378	chlamydia t
394	25	62.5	540	1	CH60_CLODI	Q9kkf0	clostridium
395	25	62.5	543	1	CH60_CLOBO	Q8kj24	clostridium
396	25	62.5	546	1	MUP3_YEAST	P38734	saccharomyc
397	25	62.5	553	1	OXAA_RALSO	Q8y3h6	ralstonia s
398	25	62.5	554	1	PEX2_PODAN	P51021	podospora a
399	25	62.5	558	1	MNT2_YEAST	P53059	saccharomyc
400	25	62.5	569	1	SYR_FUSNN	Q8rgl4	fusobacteri
401	25	62.5	575	1	HMD1_HEVBR	P29057	hevea brasi
402	25	62.5	576	1	DFA1_ANASP	Q8ynw5	anabaena sp
403	25	62.5	586	1	RPOC_EUGGR	P23580	euglena gra
404	25	62.5	599	1	DX52_HUMAN	Q9y2r4	homo sapien
405	25	62.5	609	1	GLMS_LEPIN	Q8ezq1	l glucosami
406	25	62.5	610	1	NU5M_SQUAC	Q9zz44	squalus aca
407	25	62.5	622	1	YRT1_CAEEL	Q10044	caenorhabdi
408	25	62.5	626	1	DXS_WIGBR	Q8d357	wiggleswort
409	25	62.5	671	1	Z282_HUMAN	Q9udv7	homo sapien
410	25	62.5	703	1	LAGD_LACLA	P59852	lactococcus
411	25	62.5	717	1	MCCA_MOUSE	Q99mr8	mus musculu
412	25	62.5	736	1	EF2_AERPE	Q9yc19	aeropyrum p
413	25	62.5	737	1	DPP3_HUMAN	Q9ny33	homo sapien
414	25	62.5	737	1	SIR1_MOUSE	Q923e4	mus musculu
415	25	62.5	746	1	POT1_MOUSE	Q8r2r1	mus musculu
416	25	62.5	747	1	POT1_HUMAN	Q9y6a1	homo sapien



417	25	62.5	747	1	SIR1_HUMAN	Q96eb6	homo sapien
418	25	62.5	757	1	ECR_LUCCU	O18531	lucilia cup
419	25	62.5	759	1	FXM1_RAT	P97691	rattus norv
420	25	62.5	787	1	OXAA_CHLMU	Q9pke3	chlamydia m
421	25	62.5	792	1	SYFB_CHLPN	Q9z7w0	chlamydia p
422	25	62.5	793	1	SYFB_CHLCV	Q824j8	chlamydophi
423	25	62.5	804	1	GYRB_CHLTR	O84193	chlamydia t
424	25	62.5	833	1	HMDH_AGRIP	O76819	agrotis ips
425	25	62.5	833	1	HSF_YEAST	P10961	saccharomyc
426	25	62.5	841	1	CHS1_PHYBL	P87073	phycomyces
427	25	62.5	858	1	CHS2_RHIOL	P30595	rhizopus ol
428	25	62.5	884	1	CAD1_MOUSE	P09803	mus musculu
429	25	62.5	886	1	CAD1_RAT	Q9r0t4	rattus norv
430	25	62.5	923	1	HEX_ADE40	P11819	human adeno
431	25	62.5	947	1	PODK_MAIZE	P11155	zea mays (m
432	25	62.5	951	1	HEX_ADE05	P04133	human adeno
433	25	62.5	955	1	PODK_FLABR	Q39734	flaveria br
434	25	62.5	956	1	PODK_FLAPR	Q42736	flaveria pr
435	25	62.5	967	1	HEX_ADE02	P03277	human adeno
436	25	62.5	1055	1	EPB2_HUMAN	P29323	homo sapien
437	25	62.5	1090	1	SEC5_ARATH	Q8s3u9	arabidopsis
438	25	62.5	1155	1	RPOB_BORBU	Q59191	borrelia bu
439	25	62.5	1181	1	YY02_METJA	Q60301	methanococc
440	25	62.5	1259	1	LIN1_HUMAN	P08547	homo sapien
441	25	62.5	1260	1	LIN1_NYCCO	P08548	nycticebus
442	25	62.5	1288	1	CTA3_HUMAN	Q9bz76	homo sapien
443	25	62.5	1332	1	XDH_HUMAN	P47989	homo sapien
444	25	62.5	1342	1	RPOB_BUCAI	P57146	buchnera ap
445	25	62.5	1342	1	RPOB_BUCAP	P41184	buchnera ap
446	25	62.5	1342	1	RPOB_ECOLI	P00575	escherichia
447	25	62.5	1342	1	RPOB_SALTY	P06173	salmonella
448	25	62.5	1392	1	RPOB_NEIMA	P57009	neisseria m
449	25	62.5	1392	1	RPOB_NEIMB	Q59622	neisseria m
450	25	62.5	1458	1	PHLX_RABIT	Q05017	oryctolagus
451	25	62.5	1490	1	CRK7_HUMAN	Q9nyv4	homo sapien
452	25	62.5	1585	1	YQBO_BACSU	P45931	bacillus su
453	25	62.5	1785	1	GLS3_YEAST	Q04952	saccharomyc
454	25	62.5	1786	1	YCF1_ARATH	P56785	arabidopsis
455	25	62.5	1932	1	FAB1_SCHPO	O59722	schizosacch
456	25	62.5	2021	1	OMPA_RICCN	Q52657	rickettsia
457	25	62.5	2249	1	OMPA_RICRI	P15921	rickettsia
458	25	62.5	2314	1	AKA6_RAT	Q9wvc7	rattus norv
459	25	62.5	4349	1	DYHC_FUSSO	P78716	fusarium so
460	25	62.5	4367	1	DYHC_NEUCR	P45443	neurospora
461	25	62.5	4499	1	DYHA_CHLRE	Q39610	chlamydomon
462	25	62.5	4568	1	DYHB_CHLRE	Q39565	chlamydomon
463	25	62.5	4594	1	DYHC_HUMAN	Q14204	homo sapien
464	25	62.5	4644	1	DYHC_MOUSE	Q9jhu4	mus musculu
465	25	62.5	4644	1	DYHC_RAT	P38650	rattus norv
466	25	62.5	7073	1	R1AB_CVHSA	P59641	h replicase
467	24	60.0	74	1	Y131_MYCGE	P47377	mycoplasma
468	24	60.0	76	1	CRP1_HUMAN	P50238	homo sapien
469	24	60.0	76	1	CRP1_MOUSE	P04006	mus musculu
470	24	60.0	80	1	AFP1_ARATH	P30224	arabidopsis
471	24	60.0	81	1	Y334_ARCFU	O29913	archaeoglob
472	24	60.0	84	1	YRJ7_CAEEL	Q09412	caenorhabdi
473	24	60.0	85	1	SIX2_MESMA	Q8i0k7	mesobuthus

474	24	60.0	86	1	RL23_METJA	P54016	methanococc
475	24	60.0	86	1	RL23_METVA	P10143	methanococc
476	24	60.0	94	1	Y753_TREPA	O83734	treponema p
477	24	60.0	95	1	MYLE_HUMAN	O95424	homo sapien
478	24	60.0	95	1	MYLE_MOUSE	Q9wuq7	mus musculu
479	24	60.0	100	1	PRIB_VIBVU	Q8dc15	vibrio vuln
480	24	60.0	100	1	YLXQ_BACSU	P32729	bacillus su
481	24	60.0	110	1	RBS_THINE	P45686	thiobacillu
482	24	60.0	112	1	Y145_ARCFU	O30092	archaeoglob
483	24	60.0	113	1	RBS_SYNPX	P96487	synechococc
484	24	60.0	121	1	RBFA_MYCGA	Q7nbz3	mycoplasma
485	24	60.0	123	1	RL7A_METTH	O26355	methanobact
486	24	60.0	124	1	RBFA_BUCAP	Q8k9h2	buchnera ap
487	24	60.0	135	1	NUSB_CLOPE	Q8xjd6	clostridium
488	24	60.0	138	1	YA81_METJA	Q58481	methanococc
489	24	60.0	143	1	Y771_METTH	O26865	methanobact
490	24	60.0	144	1	YE08_ARCFU	O28864	archaeoglob
491	24	60.0	147	1	YXXD_BACSU	P40737	bacillus su
492	24	60.0	152	1	SODC_ZANAE	O65174	zantedeschi
493	24	60.0	158	1	PLGF_RAT	Q63434	rattus norv
494	24	60.0	162	1	COAT_SHMV	P03581	sunh-hemp m
495	24	60.0	171	1	IL26_HUMAN	Q9nph9	homo sapien
496	24	60.0	175	1	CRG2_XENLA	Q91724	xenopus lae
497	24	60.0	175	1	CRG3_XENLA	P55940	xenopus lae
498	24	60.0	176	1	NREFX_HAEIN	P44943	haemophilus
499	24	60.0	177	1	MLF2_MALFU	P56577	malassezia
500	24	60.0	180	1	YBYK_CAEEL	P91127	caenorhabdi
501	24	60.0	191	1	YAF8_CAEEL	P52879	caenorhabdi
502	24	60.0	193	1	IL18_HUMAN	Q14116	homo sapien
503	24	60.0	196	1	HIR5_HUMAN	Q9ums0	homo sapien
504	24	60.0	196	1	Y449_BUCAI	P57524	buchnera ap
505	24	60.0	199	1	HIR5_MOUSE	Q9qz23	mus musculu
506	24	60.0	200	1	RECR_CHLPN	Q9z8n9	chlamydia p
507	24	60.0	201	1	AlAG_RABIT	P25227	oryctolagus
508	24	60.0	201	1	RYH1_SCHPO	P17608	schizosacch
509	24	60.0	203	1	CHPE_PSEAE	O87005	pseudomonas
510	24	60.0	205	1	RAB6_CAEEL	P34213	caenorhabdi
511	24	60.0	208	1	RB6A_HUMAN	P20340	homo sapien
512	24	60.0	208	1	RB6A_MOUSE	P35279	mus musculu
513	24	60.0	208	1	RB6B_HUMAN	Q9nrw1	homo sapien
514	24	60.0	210	1	ACUA_BACSU	P39065	bacillus su
515	24	60.0	211	1	YCY5_ASTLO	P14758	astasia lon
516	24	60.0	217	1	MSA1_CAUCR	Q9a9i6	caulobacter
517	24	60.0	217	1	MSA1_RHIME	Q92sy7	rhizobium m
518	24	60.0	222	1	Y052_METJA	Q60359	methanococc
519	24	60.0	224	1	TRMB_CHLMU	Q9pl91	chlamydia m
520	24	60.0	224	1	VLT3_VACCV	P07609	vaccinia vi
521	24	60.0	224	1	Y681_CHLPN	Q9z7m3	chlamydia p
522	24	60.0	225	1	VLT3_FOWPV	Q9j566	fowlpox vir
523	24	60.0	229	1	YPDP_BACSU	P54163	bacillus su
524	24	60.0	233	1	MTRP_HUMAN	Q15012	homo sapien
525	24	60.0	236	1	MTAP_SULSO	P50389	sulfolobus
526	24	60.0	238	1	RK5_ODOSI	P49547	odontella s
527	24	60.0	239	1	ATE_CAMJE	Q9pnq6	campylobact
528	24	60.0	242	1	RL1_AQUAE	O67759	aquifex aeo
529	24	60.0	244	1	YH38_HAEIN	P44302	haemophilus
530	24	60.0	246	1	NPD_THEMEA	Q9wyw0	thermotoga

531	24	60.0	249	1	UPPS_BUCAP	Q8k9s6	buchnera ap
532	24	60.0	251	1	YQ33_VIBPA	Q87lh9	vibrio para
533	24	60.0	252	1	YC85_VIBCH	Q9ksh1	vibrio chol
534	24	60.0	252	1	YE86_VIBVU	Q8dcd5	vibrio vuln
535	24	60.0	252	1	YS97_VIBVY	Q7mhh4	vibrio vuln
536	24	60.0	254	1	RB6C_HUMAN	Q9h0n0	homo sapien
537	24	60.0	255	1	TATC_AZOCH	P54085	azotobacter
538	24	60.0	256	1	TATC_HAEIN	P44560	haemophilus
539	24	60.0	257	1	6PGL_SCHPO	O74455	schizosacch
540	24	60.0	261	1	YRBE_HAEIN	P45030	haemophilus
541	24	60.0	264	1	ME24_SCHPO	Q96ws1	schizosacch
542	24	60.0	266	1	FLIP_CAUCR	Q45980	caulobacter
543	24	60.0	273	1	LACG_AGRRD	P29824	agrobacteri
544	24	60.0	276	1	YE36_AQUAE	O67427	aquifex aeo
545	24	60.0	277	1	DMA_VIBCH	Q08318	vibrio chol
546	24	60.0	280	1	YE78_STRP3	Q8k657	streptococc
547	24	60.0	280	1	YG98_STRPY	Q99yh7	streptococc
548	24	60.0	280	1	YH09_STRP8	Q8nzw1	streptococc
549	24	60.0	283	1	YS95_MYCTU	Q10816	mycobacteri
550	24	60.0	284	1	LEP_PSEFL	P26844	pseudomonas
551	24	60.0	286	1	DAFF_OCEIH	Q8enx2	oceanobacil
552	24	60.0	291	1	MSAB_TREPA	O83641	t peptide m
553	24	60.0	292	1	XTHA_PHAAN	Q41638	phaseolus a
554	24	60.0	293	1	RPC6_DROME	Q9vd25	drosophila
555	24	60.0	308	1	NOD1_RHIME	P03031	rhizobium m
556	24	60.0	309	1	ZIP2_HUMAN	Q9np94	homo sapien
557	24	60.0	311	1	CYF_SYNEL	Q9x9t1	synechococc
558	24	60.0	312	1	MAB1_STRPN	P35593	s peptide m
559	24	60.0	314	1	TRUB_VIBPA	Q87m04	vibrio para
560	24	60.0	319	1	CDK4_XENLA	Q91727	xenopus lae
561	24	60.0	328	1	HEM2_BACHD	Q9k8g2	bacillus ha
562	24	60.0	328	1	YG19_YEAST	P53208	saccharomyc
563	24	60.0	329	1	EBA3_FLAME	P36913	flavobacter
564	24	60.0	329	1	YHO3_YEAST	P14693	saccharomyc
565	24	60.0	330	1	SYW_DEIRA	Q9rwv7	deinococcus
566	24	60.0	336	1	RSMC_BUCAP	Q8k9l5	buchnera ap
567	24	60.0	340	1	ARGC_STRMU	P59311	streptococc
568	24	60.0	341	1	ARGC_STAEP	Q8cp35	staphylococ
569	24	60.0	341	1	LPXK_RHILO	P58185	rhizobium l
570	24	60.0	341	1	RTCA_PSEAE	Q9hvj9	pseudomonas
571	24	60.0	344	1	CHEB_THEMA	Q9wyn9	thermotoga
572	24	60.0	344	1	NTRI_HUMAN	Q9p121	homo sapien
573	24	60.0	345	1	ANX1_HORSE	Q8hzm6	equus cabal
574	24	60.0	346	1	ANX1_PIG	P19619	sus scrofa
575	24	60.0	348	1	YK51_YEAST	P36152	saccharomyc
576	24	60.0	358	1	COBT_CORGL	Q8nnj8	corynebacte
577	24	60.0	360	1	MRAY_HAEDU	Q7vp57	haemophilus
578	24	60.0	365	1	LMC1_HUMAN	Q9nzu5	homo sapien
579	24	60.0	367	1	RCL1_YEAST	Q08096	saccharomyc
580	24	60.0	373	1	ACT_CHOCR	P53499	chondrus cr
581	24	60.0	374	1	MTB1_BACBR	P34905	bacillus br
582	24	60.0	375	1	GP27_HUMAN	Q9ns67	homo sapien
583	24	60.0	377	1	GP27_RAT	Q9jjh3	rattus norv
584	24	60.0	379	1	GP27_MOUSE	O54897	mus musculu
585	24	60.0	384	1	NRFF_HAEIN	P44942	haemophilus
586	24	60.0	386	1	YGIC_ECOLI	P24196	escherichia
587	24	60.0	388	1	YC09_KLEPN	Q48455	klebsiella

588	24	60.0	389	1	SIR2_MOUSE	Q8vdq8	mus musculu
589	24	60.0	395	1	PGK_BRUSU	Q8fyx8	brucella su
590	24	60.0	396	1	PGK_BRUAB	Q91560	brucella ab
591	24	60.0	396	1	PGK_BRUME	Q8yiy0	brucella me
592	24	60.0	397	1	ACK2_VIBCH	Q9kmt5	vibrio chol
593	24	60.0	397	1	ACK2_VIBPA	Q87ij5	vibrio para
594	24	60.0	397	1	ACK2_VIBVU	Q8d7k4	vibrio vuln
595	24	60.0	404	1	Y4XM_RHISN	P55705	rhizobium s
596	24	60.0	405	1	ASSY_RHIME	Q92173	rhizobium m
597	24	60.0	405	1	SYX_BORBU	O51343	borrelia bu
598	24	60.0	406	1	ASSY_BRUME	Q8yek8	brucella me
599	24	60.0	406	1	ASSY_BRUSU	Q8g376	brucella su
600	24	60.0	406	1	KDC2_HUMAN	Q9y2u9	homo sapien
601	24	60.0	409	1	RHAG_HUMAN	Q02094	homo sapien
602	24	60.0	412	1	ODP2_RICCN	Q92hk7	rickettsia
603	24	60.0	413	1	APGM_SULTO	Q975p3	sulfolobus
604	24	60.0	413	1	Y286_BUCAP	Q8k9n7	buchnera ap
605	24	60.0	417	1	HFLK_BUCBP	Q89a39	buchnera ap
606	24	60.0	420	1	SYH_MYCPU	Q98qm8	mycoplasma
607	24	60.0	422	1	RA52_CHICK	P39022	gallus gall
608	24	60.0	425	1	RAFB_ECOLI	P16552	escherichia
609	24	60.0	429	1	PUR2_VIBCH	Q9kv81	vibrio chol
610	24	60.0	431	1	UL78_HCMVA	P16751	human cytom
611	24	60.0	432	1	GSAB_LISMO	Q8y6j9	listeria mo
612	24	60.0	433	1	LGMN_HUMAN	Q99538	homo sapien
613	24	60.0	435	1	LGMN_MOUSE	O89017	mus musculu
614	24	60.0	435	1	LGMN_RAT	Q9r0j8	rattus norv
615	24	60.0	436	1	VID2_AGRRH	P13462	agrobacteri
616	24	60.0	439	1	Y426_BUCAP	Q8k9c2	buchnera ap
617	24	60.0	444	1	CYB_RHOSH	Q02761	rhodobacter
618	24	60.0	447	1	NISK_LACLA	P42707	lactococcus
619	24	60.0	448	1	DHE4_HELPJ	Q9zkd8	helicobacte
620	24	60.0	448	1	DHE4_HELPY	P55990	helicobacte
621	24	60.0	452	1	MENE_HAEIN	P44565	haemophilus
622	24	60.0	456	1	GLMU_ECOLI	P17114	escherichia
623	24	60.0	457	1	HEMN_ECOLI	P32131	escherichia
624	24	60.0	457	1	HEMN_SALTY	P37129	salmonella
625	24	60.0	457	1	MURC_THEMEA	Q9wy73	thermotoga
626	24	60.0	458	1	ARLY_HAEDU	Q7vme5	haemophilus
627	24	60.0	459	1	HMT2_SCHPO	O94284	schizosacch
628	24	60.0	459	1	UTMP_BOVIN	P46201	bos taurus
629	24	60.0	459	1	UUP2_HAEIN	P45167	haemophilus
630	24	60.0	460	1	PGMU_NEIGO	P40390	neisseria g
631	24	60.0	460	1	PGMU_NEIMA	P57002	neisseria m
632	24	60.0	460	1	PGMU_NEIMB	P40391	neisseria m
633	24	60.0	461	1	TM15_PIG	Q9tsw0	sus scrofa
634	24	60.0	465	1	CG12_CANAL	P43062	candida alb
635	24	60.0	465	1	TM15_HUMAN	Q9c019	homo sapien
636	24	60.0	465	1	TM15_PANTR	Q7yr33	pan troglod
637	24	60.0	468	1	UXAC_BACTN	Q8a9j2	bacteroides
638	24	60.0	470	1	RBL_PROHO	P27568	prochloroth
639	24	60.0	470	1	RBL_SYNP2	Q44176	synechococc
640	24	60.0	471	1	MM13_BOVIN	O77656	bos taurus
641	24	60.0	471	1	YON2_YEAST	Q08271	saccharomyc
642	24	60.0	472	1	MM13_MOUSE	P33435	mus musculu
643	24	60.0	472	1	RBL_SYNP6	P00880	synechococc
644	24	60.0	476	1	MM10_HUMAN	P09238	homo sapien

645	24	60.0	478	1	RBL_NEUMU	P19163	neurachne m
646	24	60.0	478	1	RBL_NEUTE	P19164	neurachne t
647	24	60.0	480	1	DDC_CAVPO	P22781	cavia porce
648	24	60.0	480	1	NR61_HUMAN	Q15406	homo sapien
649	24	60.0	481	1	LBP_HUMAN	P18428	homo sapien
650	24	60.0	484	1	YA62_METJA	Q58462	methanococc
651	24	60.0	486	1	COXW_YEAST	P40086	saccharomyc
652	24	60.0	486	1	DDC_PIG	P80041	sus scrofa
653	24	60.0	487	1	DDC_BOVIN	P27718	bos taurus
654	24	60.0	493	1	FUT3_ARATH	Q9ca71	arabidopsis
655	24	60.0	495	1	NR61_MOUSE	Q64249	mus musculu
656	24	60.0	500	1	LIPE_HUMAN	Q9y5x9	homo sapien
657	24	60.0	502	1	C72J_ARATH	Q9ltm4	arabidopsis
658	24	60.0	502	1	C72K_ARATH	Q9ltm3	arabidopsis
659	24	60.0	504	1	GUNW_ERWCA	Q59395	erwinia car
660	24	60.0	506	1	MATK_CALVU	O47143	calluna vul
661	24	60.0	509	1	APR1_ORYSA	Q42456	oryza sativ
662	24	60.0	514	1	CLS2_BACAA	Q8ltr2	bacillus an
663	24	60.0	514	1	CLS2_BACCR	Q8lgk4	bacillus ce
664	24	60.0	514	1	QACA_STAAM	P23215	staphylococ
665	24	60.0	517	1	SEST_CAEEL	Q9n4d6	caenorhabdi
666	24	60.0	520	1	LAC1_TRAHI	Q02497	trametes hi
667	24	60.0	520	1	LAC1_TRAVI	Q99044	trametes vi
668	24	60.0	520	1	MM28_HUMAN	Q9h239	homo sapien
669	24	60.0	529	1	LAC1_PLEOS	Q12729	pleurotus o
670	24	60.0	529	1	YQP4_CAEEL	Q09531	caenorhabdi
671	24	60.0	530	1	HYAL_LYTVA	O96530	lytechinus
672	24	60.0	530	1	NTG2_HUMAN	Q96cw9	homo sapien
673	24	60.0	533	1	LAC2_PLEOS	Q12739	pleurotus o
674	24	60.0	541	1	UL21_VZVD	P09289	varicella-z
675	24	60.0	545	1	4CL1_SOLTU	P31684	solanum tub
676	24	60.0	545	1	4CL2_SOLTU	P31685	solanum tub
677	24	60.0	547	1	4CL1_TOBAC	O24145	nicotiana t
678	24	60.0	548	1	GIP2_YEAST	P40036	saccharomyc
679	24	60.0	553	1	G6PI_MYCTU	P77895	mycobacteri
680	24	60.0	554	1	NBL4_MOUSE	P52963	mus musculu
681	24	60.0	562	1	SYR_THETN	Q8r786	thermoanaer
682	24	60.0	563	1	YHJW_ECOLI	P37661	escherichia
683	24	60.0	563	1	YHJW_SALTY	P43666	salmonella
684	24	60.0	565	1	YD08_MYCPN	P75472	mycoplasma
685	24	60.0	567	1	Y391_HUMAN	O15091	homo sapien
686	24	60.0	568	1	FTSI_BUCAP	O85297	buchnera ap
687	24	60.0	569	1	PUR1_SOYBN	P52418	glycine max
688	24	60.0	570	1	GRAU_DROME	Q9u405	drosophila
689	24	60.0	573	1	SECD_MYCTU	Q50634	mycobacteri
690	24	60.0	575	1	SMF1_YEAST	P38925	saccharomyc
691	24	60.0	576	1	GPD1_ARATH	Q43727	arabidopsis
692	24	60.0	578	1	YQD6_CAEEL	Q09266	caenorhabdi
693	24	60.0	582	1	MM14_MOUSE	P53690	mus musculu
694	24	60.0	582	1	MM14_RAT	Q10739	rattus norv
695	24	60.0	586	1	HMD3_HEVBR	Q00583	hevea brasi
696	24	60.0	589	1	NTG2_MOUSE	Q8r4f1	mus musculu
697	24	60.0	589	1	SYD_XYLFA	Q9pcc5	xylella fas
698	24	60.0	589	1	SYD_XYLFT	Q87cw0	xylella fas
699	24	60.0	591	1	G6PI_PLAFA	P18240	plasmodium
700	24	60.0	594	1	OXAA_RHIME	Q92sf5	rhizobium m
701	24	60.0	598	1	DX52_MOUSE	Q8k301	mus musculu

702	24	60.0	598	1	DX52_RAT	Q99pt0	rattus norv
703	24	60.0	598	1	NBL4_HUMAN	Q9hcs5	homo sapien
704	24	60.0	610	1	VG17_BPT4	P17312	bacteriopha
705	24	60.0	611	1	RAL2_SCHPO	P15258	schizosacch
706	24	60.0	619	1	CHAC_MACFA	Q9bgz0	macaca fasc
707	24	60.0	619	1	NBL4_BRARE	O57457	brachydanio
708	24	60.0	623	1	HTPG_CLOPE	Q8xnc2	clostridium
709	24	60.0	623	1	YD54_MYCTU	Q11024	mycobacteri
710	24	60.0	625	1	HTPG_BACHD	Q9ke51	bacillus ha
711	24	60.0	626	1	HTPG_BACSU	P46208	bacillus su
712	24	60.0	629	1	MYCT_HUMAN	Q96qe2	homo sapien
713	24	60.0	631	1	PTBA_ERWCH	P26207	erwinia chr
714	24	60.0	633	1	Y147_HAEIN	P44543	haemophilus
715	24	60.0	635	1	UUP_ECOLI	P43672	escherichia
716	24	60.0	639	1	GLGB_BUTFI	P30539	butyrivibri
717	24	60.0	647	1	UUP1_HAEIN	Q57242	haemophilus
718	24	60.0	652	1	EDE1_MOUSE	Q925u4	mus musculu
719	24	60.0	652	1	T3MO_SALTY	P40814	salmonella
720	24	60.0	656	1	UVRB_CHLCV	Q821h8	chlamydophi
721	24	60.0	657	1	UVRB_CHLPN	Q9z7a5	chlamydia p
722	24	60.0	657	1	UVRB_MYCPU	P58174	mycoplasma
723	24	60.0	660	1	MM02_HUMAN	P08253	homo sapien
724	24	60.0	660	1	SYM_BACCR	Q81ja8	bacillus ce
725	24	60.0	660	1	UVB2_MYCPU	Q9zb21	mycoplasma
726	24	60.0	661	1	YN19_YEAST	P53835	saccharomyc
727	24	60.0	662	1	MM02_MOUSE	P33434	mus musculu
728	24	60.0	662	1	MM02_RABIT	P50757	oryctolagus
729	24	60.0	662	1	MM02_RAT	P33436	rattus norv
730	24	60.0	663	1	MM02_CHICK	Q90611	gallus gall
731	24	60.0	666	1	NEP1_THEVU	Q60053	thermoactin
732	24	60.0	672	1	NXF1_DROME	Q9ulh9	drosophila
733	24	60.0	675	1	NU5M_ACACA	Q37372	acanthamoeb
734	24	60.0	676	1	UVRB_CHLMU	Q9pjf4	chlamydia m
735	24	60.0	678	1	PTR3_YEAST	P43606	saccharomyc
736	24	60.0	679	1	FTSH_BACPF	P94304	bacillus ps
737	24	60.0	683	1	YJCQ_ECOLI	P32715	escherichia
738	24	60.0	685	1	CRPI_PERAM	Q25641	periplaneta
739	24	60.0	694	1	LCF4_YEAST	P47912	saccharomyc
740	24	60.0	697	1	SIM_DROME	P05709	drosophila
741	24	60.0	700	1	MR11_CHICK	Q9iam7	gallus gall
742	24	60.0	703	1	ARYB_MANSE	P14297	manduca sex
743	24	60.0	705	1	GLGB_DEIRA	Q9rtb7	deinococcus
744	24	60.0	711	1	Z336_HUMAN	Q9h116	homo sapien
745	24	60.0	722	1	BPL1_MOUSE	Q920n2	m biotin--p
746	24	60.0	724	1	PALY_USTMA	Q96v77	ustilago ma
747	24	60.0	725	1	FTKL_ANASP	Q8zs46	anabaena sp
748	24	60.0	726	1	BPL1_HUMAN	P50747	h biotin--p
749	24	60.0	730	1	PL01_CHICK	P24802	gallus gall
750	24	60.0	732	1	KMHB_DICDI	P90648	dictyosteli
751	24	60.0	743	1	PURL_RHIME	Q92ph7	rhizobium m
752	24	60.0	752	1	UGS3_PEA	Q43093	pisum sativ
753	24	60.0	769	1	GYRB_CAMJE	O87667	campylobact
754	24	60.0	775	1	PURL_AGRT5	Q8ueb0	agrobacteri
755	24	60.0	775	1	RRPO_TCV	P17460	turnip crin
756	24	60.0	778	1	EXP7_STRPN	P35597	streptococc
757	24	60.0	788	1	UGS3_SOLTU	Q43847	solanum tub
758	24	60.0	802	1	PAC_ARTVI	P31956	arthrobacte

759	24	60.0	802	1	PAC_BACME	Q60136	bacillus me
760	24	60.0	809	1	QUIA_ACICA	Q59086	acinetobact
761	24	60.0	811	1	FTSK_PSEAE	Q9i0m3	pseudomonas
762	24	60.0	828	1	V2A_TAV	P29035	tomato aspe
763	24	60.0	837	1	YL34_YEAST	Q07844	saccharomyc
764	24	60.0	852	1	R134_ARATH	Q38834	arabidopsis
765	24	60.0	856	1	POLG_PVYC	P22601	potato viru
766	24	60.0	856	1	POLG_PVYO	P22602	potato viru
767	24	60.0	875	1	ZIP1_YEAST	P31111	saccharomyc
768	24	60.0	880	1	TRK2_SCHPO	Q10065	schizosacch
769	24	60.0	885	1	CHS3_EXODE	P30602	exophiala d
770	24	60.0	889	1	CHSC_ASPFU	Q92197	aspergillus
771	24	60.0	900	1	GGAB_BACSU	P46918	bacillus su
772	24	60.0	901	1	PODK_TREPA	O83728	treponema p
773	24	60.0	911	1	CHSG_ASPFU	P54267	aspergillus
774	24	60.0	916	1	CHSB_EMENI	Q00757	emericella
775	24	60.0	919	1	YKQ5_YEAST	P36051	saccharomyc
776	24	60.0	922	1	GYRA_AERSA	P48369	aeromonas s
777	24	60.0	960	1	CHS3_NEUCR	P29070	neurospora
778	24	60.0	981	1	MCR_RAT	P22199	rattus norv
779	24	60.0	1006	1	BGAL_LACDE	P20043	lactobacill
780	24	60.0	1053	1	SLPM_BACBR	P06546	bacillus br
781	24	60.0	1057	1	SEF1_YEAST	P34228	saccharomyc
782	24	60.0	1071	1	SEF1_KLULA	P87164	kluyveromyc
783	24	60.0	1091	1	MSH3_MOUSE	P13705	mus musculu
784	24	60.0	1094	1	DPOL_GPCMV	Q69025	guinea pig
785	24	60.0	1112	1	SEN6_HUMAN	Q9gzr1	homo sapien
786	24	60.0	1121	1	CARB_MYCLE	Q9ccr2	mycobacteri
787	24	60.0	1137	1	MSH3_HUMAN	P20585	homo sapien
788	24	60.0	1147	1	NOS2_RAT	Q06518	rattus norv
789	24	60.0	1147	1	NS2D_HUMAN	O60591	homo sapien
790	24	60.0	1162	1	E2K3_DROME	Q9niv1	drosophila
791	24	60.0	1162	1	TCNA_TRYCR	P23253	trypanosoma
792	24	60.0	1169	1	EX5B_BORBU	O51578	borrelia bu
793	24	60.0	1171	1	TRI2_STRCO	Q9rkb9	streptomyce
794	24	60.0	1197	1	VGLM_RVFBV	P21401	rift valley
795	24	60.0	1206	1	VGLM_RVFBV	P03518	rift valley
796	24	60.0	1225	1	SA3_HUMAN	Q9uj98	homo sapien
797	24	60.0	1238	1	YN13_YEAST	P53840	saccharomyc
798	24	60.0	1240	1	SA3_MOUSE	O70576	mus musculu
799	24	60.0	1256	1	SA3_RAT	Q99m76	rattus norv
800	24	60.0	1362	1	PMD1_SCHPO	P36619	schizosacch
801	24	60.0	1363	1	XDH_EMENI	Q12553	emericella
802	24	60.0	1381	1	CTA1_RAT	P97846	rattus norv
803	24	60.0	1385	1	CTA1_MOUSE	O54991	mus musculu
804	24	60.0	1451	1	SPT6_YEAST	P23615	saccharomyc
805	24	60.0	1500	1	GRLF_CANFA	P83509	canis famil
806	24	60.0	1502	1	MRP6_RAT	O88269	rattus norv
807	24	60.0	1513	1	GRLF_HUMAN	Q9nry4	homo sapien
808	24	60.0	1513	1	GRLF_RAT	P81128	rattus norv
809	24	60.0	1516	1	UGG2_HUMAN	Q9nyul	homo sapien
810	24	60.0	1562	1	RPOD_CHLVU	P12465	chlorella v
811	24	60.0	1589	1	CC25_YEAST	P04821	saccharomyc
812	24	60.0	1653	1	CLH_YEAST	P22137	saccharomyc
813	24	60.0	1743	1	TAGC_DICDI	Q23868	dictyosteli
814	24	60.0	1886	1	GP21_RAT	P11654	rattus norv
815	24	60.0	1888	1	CA1E_CHICK	P32018	gallus gall

816	24	60.0	1892	1	Y835_RICCN	Q92hd6	rickettsia
817	24	60.0	2009	1	SEC7_YEAST	P11075	saccharomyc
818	24	60.0	2052	1	MY10_BOVIN	P79114	bos taurus
819	24	60.0	2058	1	MY10_HUMAN	Q9hd67	homo sapien
820	24	60.0	2150	1	SDC3_CAEEL	P34706	caenorhabdi
821	24	60.0	2209	1	KNTC_HUMAN	P50748	homo sapien
822	24	60.0	2701	1	IP3S_HUMAN	Q14571	homo sapien
823	24	60.0	2701	1	IP3S_RAT	P29995	rattus norv
824	24	60.0	2799	1	EDD_HUMAN	O95071	homo sapien
825	24	60.0	3061	1	POLG_PVYHU	Q02963	p genome po
826	24	60.0	3063	1	POLG_PVYN	P18247	p genome po
827	24	60.0	3099	1	POLG_PEMVM	O56075	p genome po
828	24	60.0	3174	1	CHAC_HUMAN	Q96rl7	homo sapien
829	24	60.0	3375	1	UN52_CAEEL	Q06561	caenorhabdi
830	24	60.0	3587	1	SRF2_BACSU	Q04747	bacillus su
831	24	60.0	3947	1	SID2_USTMA	O43103	ustilago ma
832	23.5	58.8	274	1	OXA1_STRPN	Q97ni6	streptococc
833	23.5	58.8	276	1	OXA1_STRR6	Q8dn93	streptococc
834	23.5	58.8	323	1	YF78_HAEIN	Q57287	haemophilus
835	23	57.5	29	1	SODC_OLEEU	P80740	olea europa
836	23	57.5	48	1	R332_MYCGE	Q9zb82	mycoplasma
837	23	57.5	48	1	R332_MYCPN	P56850	mycoplasma
838	23	57.5	62	1	Y085_HELPY	O24912	helicobacte
839	23	57.5	63	1	YPH2_SYNP2	P32040	synechococc
840	23	57.5	93	1	Y874_METJA	Q58284	methanococc
841	23	57.5	95	1	YN54_ARCFU	O30316	archaeoglob
842	23	57.5	100	1	DEF6_HUMAN	Q01524	homo sapien
843	23	57.5	104	1	C552_BRAJA	P24059	bradyrhizob
844	23	57.5	106	1	Y270_SYNY3	P73889	synechocyst
845	23	57.5	108	1	YY03_METJA	Q60302	methanococc
846	23	57.5	112	1	NBR2_HUMAN	O15453	homo sapien
847	23	57.5	112	1	YUW4_CAEEL	Q20263	caenorhabdi
848	23	57.5	119	1	YD52_PASMU	Q9cl88	pasteurella
849	23	57.5	121	1	CADF_STAAU	P37374	staphylococ
850	23	57.5	121	1	NU3M_CHOCR	P48910	chondrus cr
851	23	57.5	123	1	RL7A_METKA	Q8tv03	methanopyru
852	23	57.5	123	1	URE2_UREPA	Q56558	ureaplasma
853	23	57.5	123	1	URE2_UREUR	P17273	ureaplasma
854	23	57.5	124	1	FLGB_AQUAE	O67244	aquifex ae
855	23	57.5	127	1	YFFN_ECOLI	P76545	escherichia
856	23	57.5	128	1	YMJ0_YEAST	Q04501	saccharomyc
857	23	57.5	130	1	Y613_PASMU	Q9cn32	pasteurella
858	23	57.5	136	1	RET3_FUGRU	O42386	fugu rubrip
859	23	57.5	138	1	MMOB_METTR	P27356	methylosinu
860	23	57.5	139	1	NDK_CAUCR	Q9a7m2	caulobacter
861	23	57.5	140	1	NDK_METJA	Q58661	methanococc
862	23	57.5	143	1	COFI_YEAST	Q03048	saccharomyc
863	23	57.5	144	1	CNIH_HUMAN	O95406	homo sapien
864	23	57.5	144	1	CNIH_MOUSE	O35372	mus musculu
865	23	57.5	144	1	CNI_DROME	P49858	drosophila
866	23	57.5	144	1	CNI_DROVI	P52159	drosophila
867	23	57.5	144	1	R12A_ARATH	Q9s9p1	arabidopsis
868	23	57.5	145	1	Y507_RICPR	Q9zd40	rickettsia
869	23	57.5	147	1	RBFA_SHEON	Q8ehl4	shewanella
870	23	57.5	147	1	YD73_CLOTE	Q895a7	clostridium
871	23	57.5	149	1	CYB_CERNN	P82047	cervus nipp
872	23	57.5	151	1	SOD2_BRAJU	Q42612	brassica ju



873	23	57.5	151	1	SODC_NICPL	P27082	nicotiana p
874	23	57.5	153	1	CYB_CERNT	P82045	cervus nipp
875	23	57.5	154	1	CYB_CERNA	P82049	cervus nipp
876	23	57.5	154	1	CYB_CERNH	P82048	cervus nipp
877	23	57.5	154	1	CYB_CERNP	P82046	cervus nipp
878	23	57.5	155	1	FLIL_SALTY	P26417	salmonella
879	23	57.5	155	1	NUSB_VIBFI	Q8g9g3	vibrio fisc
880	23	57.5	155	1	NUSB_VIBPA	Q87ru3	vibrio para
881	23	57.5	155	1	NUSB_VIBVU	Q8dfa0	vibrio vuln
882	23	57.5	155	1	NUSB_VIBVY	Q7mn52	vibrio vuln
883	23	57.5	156	1	NUSB_VIBCH	Q9kpu5	vibrio chol
884	23	57.5	157	1	SSSP_PLAFO	P17503	plasmodium
885	23	57.5	158	1	MGSA_TRESO	P58761	treponema s
886	23	57.5	159	1	OAZ_CAEEL	Q9nhz6	caenorhabdi
887	23	57.5	160	1	CNIL_HUMAN	Q8tbel	homo sapien
888	23	57.5	161	1	VP18_NPVAC	P41481	autographa
889	23	57.5	162	1	RRAA_HAEIN	P44738	haemophilus
890	23	57.5	165	1	FTN2_HAEIN	P43708	haemophilus
891	23	57.5	165	1	RRAA_HAEDU	P59885	haemophilus
892	23	57.5	165	1	YL72_AQUAE	O67921	aquifex aeo
893	23	57.5	166	1	RRAA_PASMU	Q9clp9	pasteurella
894	23	57.5	169	1	ALLA_PSEAE	Q9i3j9	pseudomonas
895	23	57.5	169	1	YBGA_ECOLI	P24252	escherichia
896	23	57.5	172	1	RL10_RHIME	Q92qh9	rhizobium m
897	23	57.5	173	1	FABA_XYLFA	Q9pft5	xylella fas
898	23	57.5	173	1	FABA_XYLFT	Q87b86	xylella fas
899	23	57.5	173	1	YFHC_HAEIN	P44931	haemophilus
900	23	57.5	176	1	Y091_BUCAP	Q9zhe8	buchnera ap
901	23	57.5	178	1	CALC_MOUSE	Q63811	mus musculu
902	23	57.5	185	1	LEPQ_BACNA	Q57350	bacillus su
903	23	57.5	186	1	TNR5_PSEAE	P06691	pseudomonas
904	23	57.5	186	1	TNR7_ECOLI	P06692	escherichia
905	23	57.5	187	1	LPKA_METJA	Q58731	methanococc
906	23	57.5	195	1	COAG_LIMPO	P03998	limulus pol
907	23	57.5	197	1	SEQA_HAEIN	P44564	haemophilus
908	23	57.5	200	1	YKT6_YEAST	P36015	saccharomyc
909	23	57.5	203	1	RB13_HUMAN	P51153	homo sapien
910	23	57.5	203	1	YA18_METKA	Q8twl5	methanopyru
911	23	57.5	205	1	KGUA_CHLPN	Q9z961	chlamydia p
912	23	57.5	207	1	HIS2_CAMJE	Q9pm71	campylobact
913	23	57.5	208	1	FLIZ_BORBU	Q44904	borrelia bu
914	23	57.5	208	1	NPD_ACTAC	Q9zab8	actinobacil
915	23	57.5	212	1	EPD_CLUHA	P32187	clupea hare
916	23	57.5	215	1	PT18_YEAST	P25362	saccharomyc
917	23	57.5	219	1	ERD2_YEAST	P18414	saccharomyc
918	23	57.5	219	1	REP2_SCHPO	Q09824	schizosacch
919	23	57.5	219	1	YI49_AQUAE	O67701	aquifex aeo
920	23	57.5	221	1	Y013_TREPA	O83057	treponema p
921	23	57.5	222	1	Y101_MYCGE	P47347	mycoplasma
922	23	57.5	223	1	SRY_HORSE	P36389	equus cabal
923	23	57.5	226	1	YAI5_SCHPO	Q09893	schizosacch
924	23	57.5	228	1	RNC_ANASP	Q8z023	anabaena sp
925	23	57.5	228	1	VIF_OMVVS	P16902	ovine lenti
926	23	57.5	228	1	Y067_ARCFU	O30169	archaeoglob
927	23	57.5	235	1	RGSI_MOUSE	Q99pg4	mus musculu
928	23	57.5	237	1	BIOW_METJA	Q58693	methanococc
929	23	57.5	238	1	GFP_AEQVI	P42212	aequorea vi

930	23	57.5	238	1	ZPRO_MOUSE	Q9da39	mus musculu
931	23	57.5	239	1	Y4IH_RHISN	P55491	rhizobium s
932	23	57.5	240	1	RK1_CYACA	Q9tm00	cyanidium c
933	23	57.5	240	1	UNG_MYCPN	P75536	mycoplasma
934	23	57.5	242	1	Y755_AQUAE	O66957	aquifex aeo
935	23	57.5	243	1	Y358_BORBU	O51333	borrelia bu
936	23	57.5	243	1	YQEF_BACSU	P54451	bacillus su
937	23	57.5	244	1	FMCD_BACNO	P17418	bacteroides
938	23	57.5	244	1	TPIS_MYCGE	P47670	mycoplasma
939	23	57.5	245	1	UNG_MYCGE	P47343	mycoplasma
940	23	57.5	245	1	YITQ_BACSU	O06752	bacillus su
941	23	57.5	247	1	DE21_ANTMA	Q8rvl4	antirrhinum
942	23	57.5	248	1	YA14_SCHPO	Q09686	schizosacch
943	23	57.5	249	1	T4S2_HUMAN	P41732	homo sapien
944	23	57.5	249	1	T4S2_MOUSE	Q62283	mus musculu
945	23	57.5	251	1	SPEA_STRPY	P08095	streptococc
946	23	57.5	251	1	Y116_MYCGE	P47362	mycoplasma
947	23	57.5	252	1	Y476_PYRFU	Q8u3j0	pyrococcus
948	23	57.5	253	1	Y2B6_BORPE	Q7vsv6	bordetella
949	23	57.5	255	1	RL8A_YEAST	P17076	saccharomyc
950	23	57.5	255	1	RL8B_YEAST	P29453	saccharomyc
951	23	57.5	257	1	CHER_BACHD	Q9kcb8	bacillus ha
952	23	57.5	257	1	PSTB_SULSO	Q97zt9	sulfolobus
953	23	57.5	257	1	SIR3_MOUSE	Q8rl04	mus musculu
954	23	57.5	258	1	RL8_SCHPO	O13672	schizosacch
955	23	57.5	261	1	Y957_AERPE	Q9ydf6	aeropyrum p
956	23	57.5	263	1	DNAA_MYCMY	P35889	mycoplasma
957	23	57.5	265	1	NS12_ARATH	Q9lnh6	arabidopsis
958	23	57.5	265	1	TRPC_XANAC	Q8pq47	xanthomonas
959	23	57.5	265	1	TRPC_XANCP	Q8pd70	xanthomonas
960	23	57.5	267	1	COML_NEIGO	Q50985	neisseria g
961	23	57.5	267	1	COML_NEIMA	Q9jvb7	neisseria m
962	23	57.5	267	1	COML_NEIMB	Q9k0b1	neisseria m
963	23	57.5	267	1	DAPF_BACTN	Q8aab7	bacteroides
964	23	57.5	267	1	Y012_PYRFU	Q8u4r1	pyrococcus
965	23	57.5	267	1	YEBC_BACSU	O34341	bacillus su
966	23	57.5	268	1	IL1B_HORSE	Q28386	equus cabal
967	23	57.5	273	1	NPD2_BRAJA	Q89ea6	bradyrhizob
968	23	57.5	276	1	DRAT_RHORU	P14299	rhodospiril
969	23	57.5	278	1	Y751_METJA	Q58161	methanococc
970	23	57.5	280	1	GPI2_YEAST	P46961	saccharomyc
971	23	57.5	281	1	CLCX_CAEEL	Q9ngj7	caenorhabdi
972	23	57.5	281	1	NPD3_PSESM	Q882k4	pseudomonas
973	23	57.5	282	1	LRPR_STREQ	Q54087	streptococc
974	23	57.5	282	1	UPK_MYCLE	Q9cc42	mycobacteri
975	23	57.5	282	1	UPK_MYCTU	O06239	mycobacteri
976	23	57.5	283	1	DAPA_METTH	O26892	methanobact
977	23	57.5	288	1	PCD1_HUMAN	Q15116	homo sapien
978	23	57.5	288	1	ST1B_BOVIN	P41414	bos taurus
979	23	57.5	288	1	ST1B_MOUSE	P32853	mus musculu
980	23	57.5	289	1	DAPA_METJA	Q57695	methanococc
981	23	57.5	290	1	TSNX_HUMAN	Q99598	homo sapien
982	23	57.5	292	1	XTH6_ARATH	Q81f99	arabidopsis
983	23	57.5	292	1	XTH8_ARATH	Q819a9	arabidopsis
984	23	57.5	293	1	NPD_XANAC	Q8pqk3	xanthomonas
985	23	57.5	293	1	NPD_XANCP	Q8pdm9	xanthomonas
986	23	57.5	293	1	Y844_ARCFU	O29414	archaeoglob

987	23	57.5	293	1	YKUM_BACSU	O34827	bacillus su
988	23	57.5	294	1	LPXC_CAMJE	Q9piz5	campylobact
989	23	57.5	295	1	XTH_SOYBN	Q39857	glycine max
990	23	57.5	298	1	ERA_THETN	Q8rb50	thermoanaer
991	23	57.5	298	1	Y347_HELPY	O25114	helicobacte
992	23	57.5	298	1	Y485_PSEAE	Q9i633	pseudomonas
993	23	57.5	299	1	HIS1_BUCMH	Q9rq83	buchnera ap
994	23	57.5	299	1	HIS1_BUCSC	Q9rq86	buchnera ap
995	23	57.5	299	1	HIS1_YERPE	Q8zfx4	yersinia pe
996	23	57.5	299	1	NPD1_STRCO	Q9rl35	streptomyce
997	23	57.5	300	1	CH60_SYNP6	P12834	synechococc
998	23	57.5	301	1	Y376_TREPA	O83391	treponema p
999	23	57.5	302	1	RS3_HALN1	P15009	halobacteri
1000	23	57.5	303	1	NQRF_ALTHA	Q9lcj4	alteromonas

# ALIGNMENTS

## RESULT 1

### A4\_URSMA

ID A4\_URSMA STANDARD; PRT; 57 AA.  
AC Q29149;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid  
DE protein (Beta-APP) (A-beta)] (Fragment).  
GN APP.  
OS Ursus maritimus (Polar bear) (Thalarctos maritimus).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Ursus.  
OX NCBI\_TaxID=29073;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=92017079; PubMed=1656157;  
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;  
RT "Conservation of the sequence of the Alzheimer's disease amyloid  
RT peptide in dog, polar bear and five other mammals by cross-species  
RT polymerase chain reaction analysis."  
RL Brain Res. Mol. Brain Res. 10:299-305(1991).  
CC -!- FUNCTION: Functional neuronal receptor which couples to  
CC intracellular signaling pathway through the GTP-binding protein  
CC G(O) (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: Belongs to the APP family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X56128; CAA39593.1; -.

DR PIR; B60045; B60045.  
 DR HSSP; P05067; 1BA4.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; PARTIAL.  
 DR PROSITE; PS00320; A4\_INTRA; PARTIAL.  
 KW Glycoprotein; Amyloid; Neurone; Transmembrane.  
 FT NON\_TER 1 1  
 FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).  
 FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 34 57 POTENTIAL.  
 FT NON\_TER 57 57  
 SQ SEQUENCE 57 AA; 6172 MW; 84209D88EBA82DFA CRC64;

Query Match 100.0%; Score 40; DB 1; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 0.029;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||||  
 Db 21 KLVFFAED 28

## RESULT 2

### A4\_CANFA

ID A4\_CANFA STANDARD; PRT; 58 AA.  
 AC Q28280;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid  
 DE protein (Beta-APP) (A-beta)] (Fragment).  
 GN APP.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=92017079; PubMed=1656157;  
 RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;  
 RT "Conservation of the sequence of the Alzheimer's disease amyloid  
 RT peptide in dog, polar bear and five other mammals by cross-species  
 RT polymerase chain reaction analysis."  
 RL Brain Res. Mol. Brain Res. 10:299-305(1991).  
 CC -!- FUNCTION: Functional neuronal receptor which couples to  
 CC intracellular signaling pathway through the GTP-binding protein  
 CC G(O) (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Belongs to the APP family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X56125; CAA39590.1; -.  
DR HSSP; P05067; 1BA4.  
DR InterPro; IPR008155; A4\_APP.  
DR InterPro; IPR001255; Beta-APP.  
DR Pfam; PF03494; Beta-APP; 1.  
DR PROSITE; PS00319; A4\_EXTRA; PARTIAL.  
DR PROSITE; PS00320; A4\_INTRA; PARTIAL.  
KW Glycoprotein; Amyloid; Neurone; Transmembrane.  
FT NON\_TER 1 1  
FT CHAIN 7 49 BETA-AMYLOID PROTEIN (POTENTIAL).  
FT DOMAIN <1 34 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 35 58 POTENTIAL.  
FT NON\_TER 58 58  
SQ SEQUENCE 58 AA; 6285 MW; 8469D488A2E12DFA CRC64;

Query Match 100.0%; Score 40; DB 1; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.03;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
|||||||  
Db 22 KLVFFAED 29

RESULT 3

A4\_RABIT

ID A4\_RABIT STANDARD; PRT; 58 AA.  
AC Q28748;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid  
DE protein (Beta-APP) (A-beta)] (Fragment).  
GN APP.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=92017079; PubMed=1656157;  
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;  
RT "Conservation of the sequence of the Alzheimer's disease amyloid  
RT peptide in dog, polar bear and five other mammals by cross-species  
RT polymerase chain reaction analysis."  
RL Brain Res. Mol. Brain Res. 10:299-305(1991).  
CC -!- FUNCTION: Functional neuronal receptor which couples to  
CC intracellular signaling pathway through the GTP-binding protein  
CC G(O) (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: Belongs to the APP family.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X56129; CAA39594.1; -.  
DR HSSP; P05067; 1BA4.  
DR InterPro; IPR008155; A4\_APP.  
DR InterPro; IPR001255; Beta-APP.  
DR Pfam; PF03494; Beta-APP; 1.  
DR PROSITE; PS00319; A4\_EXTRA; PARTIAL.  
DR PROSITE; PS00320; A4\_INTRA; PARTIAL.  
KW Glycoprotein; Amyloid; Neurone; Transmembrane.  
FT NON\_TER 1 1  
FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).  
FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 34 57 POTENTIAL.  
FT DOMAIN 58 >58 CYTOPLASMIC (POTENTIAL).  
FT NON\_TER 58 58  
SQ SEQUENCE 58 AA; 6300 MW; F434209D88EBA82D CRC64;

Query Match 100.0%; Score 40; DB 1; Length 58;  
Best Local Similarity 100.0%; Pred. No. 0.03;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
|||  
Db 21 KLVFFAED 28

#### RESULT 4

##### A4\_SHEEP

ID A4\_SHEEP STANDARD; PRT; 58 AA.  
AC Q28757;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid  
DE protein (Beta-APP) (A-beta)] (Fragment).  
GN APP.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Heart;  
RX MEDLINE=92017079; PubMed=1656157;  
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;  
RT "Conservation of the sequence of the Alzheimer's disease amyloid  
RT peptide in dog, polar bear and five other mammals by cross-species  
RT polymerase chain reaction analysis."  
RL Brain Res. Mol. Brain Res. 10:299-305(1991).

CC -!- FUNCTION: Functional neuronal receptor which couples to  
 CC intracellular signaling pathway through the GTP-binding protein  
 CC G(O) (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Belongs to the APP family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X56130; CAA39595.1; -.  
 DR HSSP; P05067; 1BA4.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; PARTIAL.  
 DR PROSITE; PS00320; A4\_INTRA; PARTIAL.  
 KW Glycoprotein; Amyloid; Neurone; Transmembrane.  
 FT NON\_TER 1 1  
 FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).  
 FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 34 57 POTENTIAL.  
 FT DOMAIN 58 >58 CYTOPLASMIC (POTENTIAL).  
 FT NON\_TER 58 58  
 SQ SEQUENCE 58 AA; 6300 MW; F434209D88EBA82D CRC64;

Query Match 100.0%; Score 40; DB 1; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 0.03;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||||  
 Db 21 KLVFFAED 28

# RESULT 5

## A4\_BOVIN

ID A4\_BOVIN STANDARD; PRT; 59 AA.  
 AC Q28053;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid  
 DE protein (Beta-APP) (A-beta)] (Fragment).  
 GN APP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;

RX MEDLINE=92017079; PubMed=1656157;  
 RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;  
 RT "Conservation of the sequence of the Alzheimer's disease amyloid  
 RT peptide in dog, polar bear and five other mammals by cross-species  
 RT polymerase chain reaction analysis.";  
 RL Brain Res. Mol. Brain Res. 10:299-305(1991).  
 CC -!- FUNCTION: Functional neuronal receptor which couples to  
 CC intracellular signaling pathway through the GTP-binding protein  
 CC G(O) (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Belongs to the APP family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X56124; CAA39589.1; -.  
 DR EMBL; X56126; CAA39591.1; -.  
 DR HSSP; P05067; 1BA4.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; PARTIAL.  
 DR PROSITE; PS00320; A4\_INTRA; PARTIAL.  
 KW Glycoprotein; Amyloid; Neurone; Transmembrane.  
 FT NON\_TER 1 1  
 FT CHAIN 7 49 BETA-AMYLOID PROTEIN (POTENTIAL).  
 FT DOMAIN <1 34 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 35 58 POTENTIAL.  
 FT DOMAIN 59 >59 CYTOPLASMIC (POTENTIAL).  
 FT NON\_TER 59 59  
 SQ SEQUENCE 59 AA; 6414 MW; F43469D488A2E12D CRC64;

Query Match 100.0%; Score 40; DB 1; Length 59;  
 Best Local Similarity 100.0%; Pred. No. 0.03;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||||  
 Db 22 KLVFFAED 29

#### RESULT 6

##### A4\_SAISC

ID A4\_SAISC STANDARD; PRT; 751 AA.  
 AC Q95241;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Amyloid beta A4 protein precursor (APP) (Alzheimer's disease amyloid  
 DE protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha); Soluble  
 DE APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42);  
 DE Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-



DE CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)  
 DE (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-  
 DE secretase C-terminal fragment 50); C31].  
 GN APP.  
 OS *Saimiri sciureus* (Common squirrel monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.  
 OX NCBI\_TaxID=9521;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney, and Liver;  
 RX MEDLINE=96108492; PubMed=8532114;  
 RA Levy E., Amorim A., Frangione B., Walker L.C.;  
 RT "Beta-amyloid precursor protein gene in squirrel monkeys with  
 RT cerebral amyloid angiopathy."  
 RL Neurobiol. Aging 16:805-808(1995).  
 CC -!- FUNCTION: Functions as a cell surface receptor and performs  
 CC physiological functions on the surface of neurons relevant to  
 CC neurite growth, neuronal adhesion and axonogenesis. Involved in  
 CC cell mobility and transcription regulation through protein-protein  
 CC interactions (By similarity). Can promote transcription activation  
 CC through binding to APBB1/Tip60 and inhibit Notch signaling through  
 CC interaction with Numb (By similarity). Couples to apoptosis-  
 CC inducing pathways such as those mediated by G(0) and JIP (By  
 CC similarity). Inhibits G(0) alpha ATPase activity (By similarity).  
 CC Acts as a kinesin I membrane receptor, mediating the axonal  
 CC transport of beta-secretase and presenilin 1 (By similarity). May  
 CC be involved in copper homeostasis/oxidative stress through copper  
 CC ion reduction. In vitro, copper-metallated APP induces neuronal  
 CC death directly or is potentiated through Cu(II)-mediated low-  
 CC density lipoprotein oxidation (By similarity). Can regulate  
 CC neurite outgrowth through binding to components of the  
 CC extracellular matrix such as heparin and collagen I and IV (By  
 CC similarity). The splice isoforms that contain the BPTI domain  
 CC possess protease inhibitor activity (By similarity).  
 CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators  
 CC with metal-reducing activity. Bind transient metals such as  
 CC copper, zinc and iron (By similarity).  
 CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved  
 CC peptides, including C31, are potent enhancers of neuronal  
 CC apoptosis (By similarity).  
 CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several  
 CC cytoplasmic proteins, including APBB family members, the APBA  
 CC family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding  
 CC to Dab1 inhibits its serine phosphorylation (By similarity). Also  
 CC interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2  
 CC (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.  
 CC In vitro, it binds MAPT via the MT-binding domains (By  
 CC similarity). Associates with microtubules in the presence of ATP  
 CC and in a kinesin-dependent manner (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface  
 CC protein that rapidly becomes internalized via clathrin-coated  
 CC pits. During maturation, the immature APP (N-glycosylated in the  
 CC endoplasmic reticulum) moves to the Golgi complex where complete  
 CC maturation occurs (O-glycosylated and sulfated). After alpha-  
 CC secretase cleavage, soluble APP is released into the extracellular  
 CC space and the C-terminal is internalized to endosomes and

lysosomes. Some APP accumulates in secretory transport vesicles leaving the late Golgi compartment and returns to the cell surface. Gamma-CTF(59) peptide is located to both the cytoplasm and nuclei of neurons (By similarity).

CC    -!- ALTERNATIVE PRODUCTS:  
CC       Event=Alternative splicing; Named isoforms=2;  
CC       Comment=Additional isoforms seem to exist;  
CC       Name=APP770;  
CC        IsoId=Q95241-1; Sequence=Displayed;  
CC       Name=APP695;  
CC        IsoId=Q95241-2; Sequence=Not described;

CC    -!- DOMAIN: The basolateral sorting signal (BaSS) is required for sorting of membrane proteins to the basolateral surface of epithelial cells (By similarity).

CC    -!- DOMAIN: The NPXY sequence motif found in many tyrosine-phosphorylated proteins is required for the specific binding of the PID domain. However additional amino acids either N- or C-terminal to the NPXY motif are often required for complete interaction. The PID domain-containing proteins which bind APP require the YENPTY motif for full interaction. These interactions are independent of phosphorylation on the terminal tyrosine residue. The NPXY site is also involved in clathrin-mediated endocytosis (By similarity).

CC    -!- PTM: Proteolytically processed under normal cellular conditions. Cleavage by alpha-secretase or alternatively by beta-secretase leads to generation and extracellular release of soluble APP peptides, S-APP-alpha and S-APP-beta, respectively, and the retention of corresponding membrane-anchored C-terminal fragments, C83 and C99. Subsequent processing of C83 by gamma-secretase yields P3 peptides. This is the major secretory pathway and is nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated gamma-secretase processing of C99 releases the amyloid beta proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42), major components of amyloid plaques, and the cytotoxic C-terminal fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By similarity).

CC    -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis (By similarity). Cleavage at Asp-720 by either caspase-3, -8 or -9 results in the production of the neurotoxic C31 peptide and the increased production of beta-amyloid peptides (By similarity).

CC    -!- PTM: N- and O-glycosylated (By similarity).

CC    -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and serine residues is neuron-specific. Phosphorylation can affect APP processing, neuronal differentiation and interaction with other proteins (By similarity).

CC    -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and zinc, can induce histidine-bridging between beta-amyloid molecules resulting in beta-amyloid-metal aggregates (By similarity). Extracellular zinc-binding increases binding of heparin to APP and inhibits collagen-binding (By similarity).

CC    -!- SIMILARITY: Belongs to the APP family.

CC    -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

-----

CC    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; S81024; AAD14347.1; -.  
 DR HSSP; P05067; 1AAP.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR Pfam; PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 KW Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;  
 KW Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;  
 KW Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;  
 KW Proteoglycan; Amyloid; Alternative splicing.  
 FT SIGNAL 1 17 BY SIMILARITY.  
 FT CHAIN 18 751 A4 PROTEIN.  
 FT CHAIN 18 668 SOLUBLE APP-ALPHA (POTENTIAL).  
 FT CHAIN 18 652 SOLUBLE APP-BETA (POTENTIAL).  
 FT CHAIN 653 751 C99 (POTENTIAL).  
 FT CHAIN 653 694 BETA-AMYLOID PROTEIN 42 (POTENTIAL).  
 FT CHAIN 653 692 BETA-AMYLOID PROTEIN 40 (POTENTIAL).  
 FT CHAIN 669 751 C83 (POTENTIAL).  
 FT CHAIN 669 694 P3(42) (POTENTIAL).  
 FT CHAIN 669 692 P3(40) (POTENTIAL).  
 FT CHAIN 693 751 GAMMA-CTF(59) (POTENTIAL).  
 FT CHAIN 695 751 GAMMA-CTF(57) (POTENTIAL).  
 FT CHAIN 702 751 GAMMA-CTF(50) (POTENTIAL).  
 FT CHAIN 721 751 C31 (POTENTIAL).  
 FT DOMAIN 18 680 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 681 704 POTENTIAL.  
 FT DOMAIN 705 751 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 96 110 HEPARIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 181 188 ZINC-BINDING (BY SIMILARITY).  
 FT DOMAIN 291 341 BPTI/KUNITZ INHIBITOR.  
 FT DOMAIN 316 344 HEPARIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 363 428 HEPARIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 504 521 COLLAGEN-BINDING (BY SIMILARITY).  
 FT DOMAIN 713 732 INTERACTION WITH G(O)-ALPHA  
 FT (BY SIMILARITY).  
 FT DOMAIN 230 260 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 274 280 POLY-THR.  
 FT SITE 144 144 REQUIRED FOR COPPER(II) REDUCTION  
 FT (BY SIMILARITY).  
 FT ACT\_SITE 301 302 REACTIVE BOND.  
 FT SITE 652 653 CLEAVAGE (BY BETA-SECRETASE)

FT				(BY SIMILARITY).
FT	SITE	653	654	CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT	SITE	668	669	CLEAVAGE (BY ALPHA-SECRETASE)
FT				(BY SIMILARITY).
FT	SITE	685	685	INVOLVED IN FREE RADICAL PROPAGATION
FT				(BY SIMILARITY).
FT	SITE	687	687	INVOLVED IN OXIDATIVE REACTIONS
FT				(BY SIMILARITY).
FT	SITE	692	693	CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT				(BY SIMILARITY).
FT	SITE	694	695	CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT				(BY SIMILARITY).
FT	SITE	701	702	CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT				(BY SIMILARITY).
FT	SITE	705	715	BASOLATERAL SORTING SIGNAL
FT				(BY SIMILARITY).
FT	SITE	720	721	CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)
FT				(BY SIMILARITY).
FT	SITE	738	741	ENDOCYTOSIS SIGNAL.
FT	SITE	740	743	NPXY MOTIF.

Query Match 100.0%; Score 40; DB 1; Length 751;  
 Best Local Similarity 100.0%; Pred. No. 0.37;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||||  
 Db 668 KLVFFAED 675

# RESULT 7

## A4\_CAVPO

ID A4\_CAVPO STANDARD; PRT; 770 AA.  
 AC Q60495; Q60496;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease  
 DE amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);  
 DE Soluble APP-beta (S-APP-beta); CTF-alpha; CTF-beta; Beta-amyloid  
 DE protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); P3(42);  
 DE P3(40); CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-  
 DE CTF(57) (Gamma-secretase C-terminal fragment 57); C31].  
 GN APP.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Brain, and Liver;  
 RX MEDLINE=97236426; PubMed=9116031;  
 RA Beck M., Mueller D., Bigl V.;  
 RT "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and  
 RT alternative splicing."  
 RL Biochim. Biophys. Acta 1351:17-21(1997).  
 RN [2]

RP INTERACTION OF BETA-APP40 WITH APOE.  
 RX MEDLINE=98007700; PubMed=9349544;  
 RA Martel C.L., Mackic J.B., Matsubara E., Governale S., Miguel C.,  
 RA Miao W., McComb J.G., Frangione B., Ghiso J., Zlokovic B.V.;  
 RT "Isoform-specific effects of apolipoproteins E2, E3, and E4 on  
 RT cerebral capillary sequestration and blood-brain barrier transport of  
 RT circulating Alzheimer's amyloid beta.";  
 RL J. Neurochem. 69:1995-2004(1997).  
 RN [3]  
 RP PROCESSING.  
 RX MEDLINE=20084499; PubMed=10619481;  
 RA Beck M., Brueckner M.K., Holzer M., Kaap S., Pannicke T., Arendt T.,  
 RA Bigl V.;  
 RT "Guinea-pig primary cell cultures provide a model to study expression  
 RT and amyloidogenic processing of endogenous amyloid precursor  
 RT protein.";  
 RL Neuroscience 95:243-254(2000).  
 RN [4]  
 RP GAMMA-SECRETASE PROCESSING.  
 RX MEDLINE=20576391; PubMed=11035007;  
 RA Pinnix I., Musunuru U., Tun H., Sridharan A., Golde T., Eckman C.,  
 RA Ziani-Cherif C., Onstead L., Sambamurti K.;  
 RT "A novel gamma -secretase assay based on detection of the putative  
 RT C-terminal fragment-gamma of amyloid beta protein precursor.";  
 RL J. Biol. Chem. 276:481-487(2001).  
 CC -!- FUNCTION: Functions as a cell surface receptor and performs  
 CC physiological functions on the surface of neurons relevant to  
 CC neurite growth, neuronal adhesion and axonogenesis. Involved in  
 CC cell mobility and transcription regulation through protein-protein  
 CC interactions (By similarity). Can promote transcription activation  
 CC through binding to APBB1/Tip60 and inhibit Notch signaling through  
 CC interaction with Numb (By similarity). Couples to apoptosis-  
 CC inducing pathways such as those mediated by G(O) and JIP (By  
 CC similarity). Inhibits G(O) alpha ATPase activity (By similarity).  
 CC Acts as a kinesin I membrane receptor, mediating the axonal  
 CC transport of beta-secretase and presenilin 1 (By similarity). May  
 CC be involved in copper homeostasis/oxidative stress through copper  
 CC ion reduction (By similarity). In vitro, copper-metallated APP  
 CC induces neuronal death directly or is potentiated through Cu(II)-  
 CC mediated low-density lipoprotein oxidation (By similarity). Can  
 CC regulate neurite outgrowth through binding to components of the  
 CC extracellular matrix such as heparin and collagen I and IV (By  
 CC similarity). The splice isoforms that contain the BPTI domain  
 CC possess protease inhibitor activity (By similarity).  
 CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators  
 CC with metal-reducing activity. Bind transient metals such as  
 CC copper, zinc and iron. Beta-amyloid peptides bind to lipoproteins  
 CC and apolipoproteins E and J in the CSF and to HDL particles in  
 CC plasma, inhibiting metal-catalyzed oxidation of lipoproteins.  
 CC -!- FUNCTION: Appicans elicit adhesion of neural cells to the  
 CC extracellular matrix and may regulate neurite outgrowth in the  
 CC brain (By similarity).  
 CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved  
 CC peptides, including C31, are potent enhancers of neuronal  
 CC apoptosis (By similarity).  
 CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several  
 CC cytoplasmic proteins, including APBB family members; the APBA

family, MAPK8IP1, SHC1 and Numb and Dab1 (By similarity). Also interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2 (via its TPR domains), APPBP2 (via BaSS) and DDB1 (By similarity). Associates with microtubules in the presence of ATP and in a kinesin-dependent manner (By similarity). Soluble Abeta40 binds all three isoforms of APOE, in vitro and in vivo. When lipidated, ApoE3 appears to be the preferred amyloid binding isoform, while the apoE4 isoform-beta-APP40 complex is capable of being transported across the blood-brain barrier.

-!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface protein that rapidly becomes internalized via clathrin-coated pits (By similarity). During maturation, the immature APP (N-glycosylated in the endoplasmic reticulum) moves to the Golgi complex where complete maturation occurs (O-glycosylated and sulfated) (By similarity). After alpha-secretase cleavage, soluble APP is released into the extracellular space and the C-terminal is internalized to endosomes and lysosomes (By similarity). Some APP accumulates in secretory transport vesicles leaving the late Golgi compartment and returns to the cell surface (By similarity). APP sorts to the basolateral surface in epithelial cells (By similarity).

-!- ALTERNATIVE PRODUCTS:  
 Event=Alternative splicing; Named isoforms=2;  
 Comment=Additional isoforms, missing exons 7,8 and 15, seem to exist. The L-isoforms, missing exon 15, are referred to as appicans;  
 Name=APP770;  
 IsoId=Q60495-1; Sequence=Displayed;  
 Name=APP695;  
 IsoId=Q60495-2; Sequence=VSP\_007221, VSP\_007222;

-!- TISSUE SPECIFICITY: Isoform APP695 is the major isoform found in brain. The longer isoforms containing the BPTI domain are predominantly expressed in peripheral organs such as muscle and liver.

-!- INDUCTION: Increased levels during neuronal differentiation.

-!- DOMAIN: The basolateral sorting signal (BaSS) is required for sorting of membrane proteins to the basolateral surface of epithelial cells.

-!- DOMAIN: The NPXY sequence motif found in many tyrosine-phosphorylated proteins is required for the specific binding of the PID domain. However additional amino acids either N- or C-terminal to the NPXY motif are often required for complete interaction. The PID domain-containing proteins which bind APP require the YENPTY motif for full interaction. These interactions are independent of phosphorylation on the terminal tyrosine residue (By similarity). The NPXY site is also involved in clathrin-mediated endocytosis.

-!- PTM: Proteolytically processed under normal cellular conditions. Cleavage by alpha-secretase or alternatively by beta-secretase leads to generation and extracellular release of soluble APP peptides, S-APP-alpha and S-APP-beta, respectively, and the retention of corresponding membrane-anchored C-terminal fragments, CTF-alpha and CTF-beta. Subsequent processing of CTF-alpha by gamma-secretase yields P3 peptides. This is the major secretory pathway and is nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated gamma-secretase processing of CTF-beta releases the amyloid beta proteins, amyloid-beta 40 (Abeta40)

CC and amyloid-beta 42 (Abeta42), major components of amyloid  
 CC plaques, and the corresponding cytotoxic C-terminal fragments  
 CC (CTFs).  
 CC -!- PTM: Proteolytically cleaved by caspase-3 during neuronal  
 CC apoptosis (By similarity).  
 CC -!- PTM: N- and O-glycosylated. O-linkage of chondroitin sulfate to  
 CC the L-APP isoforms produces the APP proteoglycan core proteins,  
 CC the appicans (By similarity).  
 CC -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and  
 CC serine residues is neuron-specific (By similarity).  
 CC Phosphorylation can affect APP processing, neuronal  
 CC differentiation and interaction with other proteins.  
 CC -!- PTM: Extracellular binding and reduction of copper, results in a  
 CC corresponding oxidation of Cys-144 and Cys-158, and the formation  
 CC of a disulfide bond (By similarity).  
 CC -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and  
 CC zinc, can induce histidine-bridging between beta-amyloid molecules  
 CC resulting in beta-amyloid-metal aggregates.  
 CC -!- SIMILARITY: Belongs to the APP family.  
 CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X97631; CAA66230.1; -.  
 DR EMBL; X99198; CAA67589.1; -.  
 DR HSSP; P05067; 1BA4.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 KW Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;  
 KW Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;  
 KW Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;  
 KW Proteoglycan; Alternative splicing; Amyloid.  
 FT SIGNAL 1 17 BY SIMILARITY.  
 FT CHAIN 18 770 AMYLOID BETA A4 PROTEIN.  
 FT CHAIN 18 687 SOLUBLE APP-ALPHA (BY SIMILARITY).  
 FT CHAIN 18 671 SOLUBLE APP-BETA (BY SIMILARITY).  
 FT CHAIN 672 770 CTF-ALPHA (BY SIMILARITY).  
 FT CHAIN 672 713 BETA-AMYLOID PROTEIN 42 (BY SIMILARITY).  
 FT CHAIN 672 711 BETA-AMYLOID PROTEIN 40 (BY SIMILARITY).  
 FT CHAIN 688 770 CTF-BETA (BY SIMILARITY).

FT	CHAIN	688	713	P3(42) (BY SIMILARITY).
FT	CHAIN	688	711	P3(40) (BY SIMILARITY).
FT	CHAIN	712	770	GAMMA-CTF(59) (BY SIMILARITY).
FT	CHAIN	714	770	GAMMA-CTF(57) (BY SIMILARITY).

Query Match 100.0%; Score 40; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||||  
 Db 687 KLVFFAED 694

# RESULT 8

## A4\_HUMAN

ID A4\_HUMAN STANDARD; PRT; 770 AA.  
 AC P05067; P09000; P78438; Q13764; Q13778; Q13793; Q16011; Q16014;  
 AC Q16019; Q16020; Q9BT38; Q9UCA9; Q9UCB6; Q9UCC8; Q9UCD1; Q9UQ58;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease  
 DE amyloid protein) (Cerebral vascular amyloid peptide) (CVAP) (Protease  
 DE nexin-II) (PN-II) (APPI) (PreA4) [Contains: Soluble APP-alpha (S-APP-  
 DE alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42  
 DE (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42);  
 DE P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59)  
 DE (Amyloid intracellular domain 59) (AID(59)); Gamma-CTF(57) (Gamma-  
 DE secretase C-terminal fragment 57) (Amyloid intracellular domain 57)  
 DE (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50)  
 DE (Amyloid intracellular domain 50) (AID(50)); C31].  
 GN APP OR A4 OR AD1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM APP695).  
 RC TISSUE=Brain;  
 RX MEDLINE=87144572; PubMed=2881207;  
 RA Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L.,  
 RA Grzeschik K.-H., Multhaup G., Beyreuther K., Mueller-Hill B.;  
 RT "The precursor of Alzheimer's disease amyloid A4 protein resembles a  
 RT cell-surface receptor."  
 RL Nature 325:733-736(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM APP751).  
 RC TISSUE=Brain;  
 RX MEDLINE=88122639; PubMed=2893289;  
 RA Ponte P., Gonzalez-Dewhitt P., Schilling J., Miller J., Hsu D.,  
 RA Greenberg B., Davis K., Wallace W., Lieberburg I., Fuller F.,  
 RA Cordell B.;  
 RT "A new A4 amyloid mRNA contains a domain homologous to serine  
 RT proteinase inhibitors."  
 RL Nature 331:525-527(1988).  
 RN [3]



RP SEQUENCE FROM N.A. (ISOFORM APP695).  
 RX MEDLINE=89128427; PubMed=2783775;  
 RA Lemaire H.-G., Salbaum J.M., Multhaup G., Kang J., Bayney R.M.,  
 RA Unterbeck A., Beyreuther K., Mueller-Hill B.;  
 RT "The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid  
 RT is encoded by 16 exons.";  
 RL Nucleic Acids Res. 17:517-522(1989).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM APP770).  
 RX MEDLINE=90236318; PubMed=2110105;  
 RA Yoshikai S.-I., Sasaki H., Doh-Ura K., Furuya H., Sakaki Y.;  
 RT "Genomic organization of the human amyloid beta-protein precursor  
 RT gene.";  
 RL Gene 87:257-263(1990).  
 RN [5]  
 RP ERRATUM, AND REVISIONS.  
 RA Yoshikai S.-I., Sasaki H., Doh-ura K., Furuya H., Sakaki Y.;  
 RL Gene 102:291-292(1991).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM L-APP733).  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=92268136; PubMed=1587857;  
 RA Koenig G., Moenning U., Czech C., Prior R., Banati R.,  
 RA Schreiter-Gasser U., Bauer J., Masters C.L., Beyreuther K.;  
 RT "Identification and differential expression of a novel alternative  
 RT splice isoform of the beta A4 amyloid precursor protein (APP) mRNA in  
 RT leukocytes and brain microglial cells.";  
 RL J. Biol. Chem. 267:10804-10809(1992).  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM APP770).  
 RX MEDLINE=97263807; PubMed=9108164;  
 RA Hattori M., Tsukahara F., Furuhashi Y., Tanahashi H., Hirose M.,  
 RA Saito M., Tsukuni S., Sakaki Y.;  
 RT "A novel method for making nested deletions and its application for  
 RT sequencing of a 300 kb region of human APP locus.";  
 RL Nucleic Acids Res. 25:1802-1808(1997).  
 RN [8]  
 RP SEQUENCE FROM N.A. (ISOFORM APP639).  
 RC TISSUE=Brain;  
 RX MEDLINE=22744650; PubMed=12859342;  
 RA Tang K., Wang C., Shen C., Sheng S., Ravid R., Jing N.;  
 RT "Identification of a novel alternative splicing isoform of human  
 RT amyloid precursor protein gene, APP639.";  
 RL Eur. J. Neurosci. 18:102-108(2003).  
 RN [9]  
 RP SEQUENCE FROM N.A. (ISOFORM APP305).  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ussdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [10]  
 RP SEQUENCE OF 1-10 FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=89016647; PubMed=3140222;  
 RA Schon E.A., Mita S., Sadlock J., Herbert J.;  
 RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)  
 RT encodes a 95-kDa polypeptide.";  
 RL Nucleic Acids Res. 16:9351-9351(1988).  
 RN [11]  
 RP ERRATUM, AND REVISIONS.  
 RA Mita S., Sadlock J., Herbert J., Schon E.A.;  
 RL Nucleic Acids Res. 16:11402-11402(1988).  
 RN [12]  
 RP SEQUENCE OF 1-75 FROM N.A.  
 RX MEDLINE=89165870; PubMed=2538123;  
 RA La Fauci G., Lahiri D.K., Salton S.R., Robakis N.K.;  
 RT "Characterization of the 5'-end region and the first two exons of the  
 RT beta-protein precursor gene.";  
 RL Biochem. Biophys. Res. Commun. 159:297-304(1989).  
 RN [13]  
 RP SEQUENCE OF 18-50.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=87250462; PubMed=3597385;  
 RA van Nostrand W.E., Cunningham D.D.;  
 RT "Purification of protease nexin II from human fibroblasts.";  
 RL J. Biol. Chem. 262:8508-8514(1987).  
 RN [14]  
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP751).  
 RC TISSUE=Brain;  
 RX MEDLINE=89346754; PubMed=2569763;  
 RA de Sauvage F., Octave J.N.;  
 RT "A novel mRNA of the A4 amyloid precursor gene coding for a possibly  
 RT secreted protein.";  
 RL Science 245:651-653(1989).  
 RN [15]  
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP695).  
 RC TISSUE=Brain;  
 RX MEDLINE=87231971; PubMed=3035574;  
 RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;  
 RT "Molecular cloning and characterization of a cDNA encoding the  
 RT cerebrovascular and the neuritic plaque amyloid peptides.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).  
 RN [16]  
 RP SEQUENCE OF 286-366 FROM N.A.  
 RX MEDLINE=88122640; PubMed=2893290;  
 RA Tanzi R.E., McClatchey A.I., Lamperti E.D., Villa-Komaroff L.,

RA Gusella J.F., Neve R.L.;  
 RT "Protease inhibitor domain encoded by an amyloid protein precursor  
 RT mRNA associated with Alzheimer's disease.";  
 RL Nature 331:528-530(1988).  
 RN [17]  
 RP SEQUENCE OF 287-367 FROM N.A.  
 RX MEDLINE=88122641; PubMed=2893291;  
 RA Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.;  
 RT "Novel precursor of Alzheimer's disease amyloid protein shows  
 RT protease inhibitory activity.";  
 RL Nature 331:530-532(1988).  
 RN [18]  
 RP SEQUENCE OF 507-770 FROM N.A.  
 RC TISSUE=Brain cortex;  
 RX MEDLINE=88124954; PubMed=2893379;  
 RA Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,  
 RA Marotta C.A.;  
 RT "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer  
 RT disease brain: coding and noncoding regions of the fetal precursor  
 RT mRNA are expressed in the cortex.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).  
 RN [19]  
 RP SEQUENCE OF 523-555, AND COLLAGEN-BINDING DOMAIN.  
 RX MEDLINE=96139497; PubMed=8576160;  
 RA Behr D., Hesse L., Masters C.L., Multhaup G.;  
 RT "Regulation of amyloid protein precursor (APP) binding to collagen and  
 RT mapping of the binding sites on APP and collagen type I.";  
 RL J. Biol. Chem. 271:1613-1620(1996).  
 RN [20]  
 RP SEQUENCE OF 655-737 FROM N.A., AND VARIANTS AD PHE-717; AD ILE-717  
 RP AND AD GLY-717.  
 RX MEDLINE=93236601; PubMed=8476439;  
 RA Denman R.B., Rosenzwaig R., Miller D.L.;  
 RT "A system for studying the effect(s) of familial Alzheimer disease  
 RT mutations on the processing of the beta-amyloid peptide precursor.";  
 RL Biochem. Biophys. Res. Commun. 192:96-103(1993).  
 RN [21]  
 RP SEQUENCE OF 656-737 FROM N.A.  
 RX MEDLINE=89392030; PubMed=2675837;  
 RA Johnstone E.M., Chaney M.O., Moore R.E., Ward K.E., Norris F.H.,  
 RA Little S.P.;  
 RT "Alzheimer's disease amyloid peptide is encoded by two exons and shows  
 RT similarity to soybean trypsin inhibitor.";  
 RL Biochem. Biophys. Res. Commun. 163:1248-1255(1989).  
 RN [22]

Query Match 100.0%; Score 40; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 687 KLVFFAED 694

RESULT 9  
 A4\_MACFA

ID A4\_MACFA STANDARD; PRT; 770 AA.  
 AC P53601; Q95KN7;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease  
 DE amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);  
 DE Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-  
 DE APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40);  
 DE Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)  
 DE (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-  
 DE secretase C-terminal fragment 50); C31].  
 GN APP.  
 OS *Macaca fascicularis* (Crab eating macaque) (*Cynomolgus* monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; *Macaca*.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS APP695 AND APP770).  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=91273117; PubMed=1905108;  
 RA Podlisy M.B., Tolan D.R., Selkoe D.J.;  
 RT "Homology of the amyloid beta protein precursor in monkey and human  
 RT supports a primate model for beta amyloidosis in Alzheimer's  
 RT disease.";  
 RL Am. J. Pathol. 138:1423-1435(1991).  
 CC -!- FUNCTION: Functions as a cell surface receptor and performs  
 CC physiological functions on the surface of neurons relevant to  
 CC neurite growth, neuronal adhesion and axonogenesis. Involved in  
 CC cell mobility and transcription regulation through protein-protein  
 CC interactions (By similarity). Can promote transcription activation  
 CC through binding to APBB1/Tip60 and inhibit Notch signaling through  
 CC interaction with Numb (By similarity). Couples to apoptosis-  
 CC inducing pathways such as those mediated by G(O) and JIP (By  
 CC similarity). Inhibits G(O) alpha ATPase activity (By similarity).  
 CC Acts as a kinesin I membrane receptor, mediating the axonal  
 CC transport of beta-secretase and presenilin 1 (By similarity). May  
 CC be involved in copper homeostasis/oxidative stress through copper  
 CC ion reduction. In vitro, copper-metallated APP induces neuronal  
 CC death directly or is potentiated through Cu(II)-mediated low-  
 CC density lipoprotein oxidation (By similarity). Can regulate  
 CC neurite outgrowth through binding to components of the  
 CC extracellular matrix such as heparin and collagen I and IV (By  
 CC similarity). The splice isoforms that contain the BPTI domain  
 CC possess protease inhibitor activity (By similarity).  
 CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators  
 CC with metal-reducing activity. Bind transient metals such as  
 CC copper, zinc and iron (By similarity).  
 CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved  
 CC peptides, including C31, are potent enhancers of neuronal  
 CC apoptosis (By similarity).  
 CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several  
 CC cytoplasmic proteins, including APBB family members, the APBA  
 CC family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding  
 CC to Dab1 inhibits its serine phosphorylation (By similarity). Also  
 CC interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2

(via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1. In vitro, it binds MAPT via the MT-binding domains (By similarity). Associates with microtubules in the presence of ATP and in a kinesin-dependent manner (By similarity).

CC    -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface protein that rapidly becomes internalized via clathrin-coated pits. During maturation, the immature APP (N-glycosylated in the endoplasmic reticulum) moves to the Golgi complex where complete maturation occurs (O-glycosylated and sulfated). After alpha-secretase cleavage, soluble APP is released into the extracellular space and the C-terminal is internalized to endosomes and lysosomes. Some APP accumulates in secretory transport vesicles leaving the late Golgi compartment and returns to the cell surface. Gamma-CTF(59) peptide is located to both the cytoplasm and nuclei of neurons (By similarity).

CC    -!- ALTERNATIVE PRODUCTS:

CC       Event=Alternative splicing; Named isoforms=2;

CC       Comment=Additional isoforms seem to exist;

CC       Name=APP770;

CC       IsoId=P53601-1; Sequence=Displayed;

CC       Name=APP695;

CC       IsoId=P53601-2; Sequence=VSP\_000010, VSP\_000011;

CC    -!- DOMAIN: The basolateral sorting signal (BaSS) is required for sorting of membrane proteins to the basolateral surface of epithelial cells (By similarity).

CC    -!- DOMAIN: The NPXY sequence motif found in many tyrosine-phosphorylated proteins is required for the specific binding of the PID domain. However additional amino acids either N- or C-terminal to the NPXY motif are often required for complete interaction. The PID domain-containing proteins which bind APP require the YENPTY motif for full interaction. These interactions are independent of phosphorylation on the terminal tyrosine residue. The NPXY site is also involved in clathrin-mediated endocytosis (By similarity).

CC    -!- PTM: Proteolytically processed under normal cellular conditions. Cleavage by alpha-secretase or alternatively by beta-secretase leads to generation and extracellular release of soluble APP peptides, S-APP-alpha and S-APP-beta, respectively, and the retention of corresponding membrane-anchored C-terminal fragments, C83 and C99. Subsequent processing of C83 by gamma-secretase yields P3 peptides. This is the major secretory pathway and is nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated gamma-secretase processing of C99 releases the amyloid beta proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42), major components of amyloid plaques, and the cytotoxic C-terminal fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By similarity).

CC    -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9 results in the production of the neurotoxic C31 peptide and the increased production of beta-amyloid peptides (By similarity).

CC    -!- PTM: N- and O-glycosylated (By similarity).

CC    -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and serine residues is neuron-specific. Phosphorylation can affect APP processing, neuronal differentiation and interaction with other proteins (By similarity).

CC    -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and

CC zinc, can induce histidine-bridging between beta-amyloid molecules  
 CC resulting in beta-amyloid-metal aggregates (By similarity).  
 CC Extracellular zinc-binding increases binding of heparin to APP and  
 CC inhibits collagen-binding (By similarity).  
 CC -!- SIMILARITY: Belongs to the APP family.  
 CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M58727; AAA36829.1; -.  
 DR EMBL; M58726; AAA36828.1; -.  
 DR HSSP; P05067; 1AAP.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR Pfam; PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 KW Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;  
 KW Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;  
 KW Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;  
 KW Proteoglycan; Alternative splicing; Amyloid.  
 FT SIGNAL 1 17 BY SIMILARITY.  
 FT CHAIN 18 770 AMYLOID BETA A4 PROTEIN.  
 FT CHAIN 18 687 SOLUBLE APP-ALPHA (POTENTIAL).  
 FT CHAIN 18 671 SOLUBLE APP-BETA (POTENTIAL).  
 FT CHAIN 672 770 C99 (POTENTIAL).  
 FT CHAIN 672 713 BETA-AMYLOID PROTEIN 42 (POTENTIAL).  
 FT CHAIN 672 711 BETA-AMYLOID PROTEIN 40 (POTENTIAL).  
 FT CHAIN 688 770 C83 (POTENTIAL).  
 FT CHAIN 688 713 P3(42) (POTENTIAL).  
 FT CHAIN 688 711 P3(40) (POTENTIAL).  
 FT CHAIN 712 770 GAMMA-CTF(59) (POTENTIAL).  
 FT CHAIN 714 770 GAMMA-CTF(57) (POTENTIAL).  
 FT CHAIN 721 770 GAMMA-CTF(50) (POTENTIAL).  
 FT CHAIN 740 770 C31 (POTENTIAL).  
 FT DOMAIN 18 699 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 700 723 POTENTIAL.  
 FT DOMAIN 724 770 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 96 110 HEPARIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 181 188 ZINC-BINDING (BY SIMILARITY).

FT	DOMAIN	291	341	BPTI/KUNITZ INHIBITOR.
FT	DOMAIN	391	423	HEPARIN-BINDING (BY SIMILARITY).
FT	DOMAIN	491	522	HEPARIN-BINDING (BY SIMILARITY).
FT	DOMAIN	523	540	COLLAGEN-BINDING (BY SIMILARITY).
FT	DOMAIN	732	751	INTERACTION WITH G(O)-ALPHA
FT				(BY SIMILARITY).
FT	DOMAIN	230	260	ASP/GLU-RICH (ACIDIC).
FT	DOMAIN	274	280	POLY-THR.
FT	SITE	144	144	REQUIRED FOR COPPER(II) REDUCTION
FT				(BY SIMILARITY).
FT	ACT_SITE	301	302	REACTIVE BOND (BY SIMILARITY).
FT	SITE	671	672	CLEAVAGE (BY BETA-SECRETASE)
FT				(BY SIMILARITY).
FT	SITE	672	673	CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT	SITE	687	688	CLEAVAGE (BY ALPHA-SECRETASE)
FT				(BY SIMILARITY).
FT	SITE	704	704	IMPLICATED IN FREE RADICAL PROPAGATION
FT				(BY SIMILARITY).
FT	SITE	706	706	INVOLVED IN OXIDATIVE REACTIONS
FT				(BY SIMILARITY).
FT	SITE	711	712	CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT				(BY SIMILARITY).
FT	SITE	713	714	CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT				(BY SIMILARITY).
FT	SITE	720	721	CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT				(BY SIMILARITY).
FT	SITE	724	734	BASOLATERAL SORTING SIGNAL
FT				(BY SIMILARITY).
FT	SITE	739	740	CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)

Query Match 100.0%; Score 40; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 687 KLVFFAED 694

# RESULT 10

## A4\_MOUSE

ID A4\_MOUSE STANDARD; PRT; 770 AA.  
 AC P12023; P97487; P97942; Q99K32;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease  
 DE amyloid protein homolog) (Amyloidogenic glycoprotein) (AG) [Contains:  
 DE Soluble APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99  
 DE (APP-C99); Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein  
 DE 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase  
 DE C-terminal fragment 59) (Amyloid intracellular domain 59) (AID(59))  
 DE (APP-C59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57)  
 DE (Amyloid intracellular domain 57) (AID(57)) (APP-C57); Gamma-CTF(50)  
 DE (Gamma-secretase C-terminal fragment 50) (Amyloid intracellular domain  
 DE 50) (AID(50)); C31].  
 GN APP.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM APP695).  
 RC TISSUE=Brain;  
 RX MEDLINE=88106489; PubMed=3322280;  
 RA Yamada T., Sasaki H., Furuya H., Miyata T., Goto I., Sakaki Y.;  
 RT "Complementary DNA for the mouse homolog of the human amyloid beta  
 RT protein precursor.";  
 RL Biochem. Biophys. Res. Commun. 149:665-671(1987).  
 RN [2]  
 RP REVISIONS.  
 RA Yamada T.;  
 RL Submitted (MAR-1988) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM APP695).  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=92096458; PubMed=1756177;  
 RA de Strooper B., van Leuven F., van den Berghe H.;  
 RT "The amyloid beta protein precursor or proteinase nexin II from mouse  
 RT is closer related to its human homolog than previously reported.";  
 RL Biochim. Biophys. Acta 1129:141-143(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM APP695).  
 RC STRAIN=SAMP8; TISSUE=Hippocampus;  
 RX MEDLINE=21130647; PubMed=11235921;  
 RA Kumar V.B., Vyas K., Franko M., Choudhary V., Buddhiraju C.,  
 RA Alvarez J., Morley J.E.;  
 RT "Molecular cloning, expression, and regulation of hippocampal amyloid  
 RT precursor protein of senescence accelerated mouse (SAMP8).";  
 RL Biochem. Cell Biol. 79:57-67(2001).  
 RN [5]  
 RP SEQUENCE OF 1-19 FROM N.A.  
 RX MEDLINE=92209998; PubMed=1555768;  
 RA Izumi R., Yamada T., Yoshikai S.I., Sasaki H., Hattori M.,  
 RA Sakai Y.;  
 RT "Positive and negative regulatory elements for the expression of the  
 RT Alzheimer's disease amyloid precursor-encoding gene in mouse.";  
 RL Gene 112:189-195(1992).  
 RN [6]  
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP770).  
 RC TISSUE=Breast tumor;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uzdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,



RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [7]  
 RP SEQUENCE OF 281-380 FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Brain, and Kidney;  
 RX MEDLINE=89149813; PubMed=2493250;  
 RA Yamada T., Sasaki H., Dohura K., Goto I., Sakaki Y.;  
 RT "Structure and expression of the alternatively-spliced forms of mRNA  
 RT for the mouse homolog of Alzheimer's disease amyloid beta protein  
 RT precursor.";  
 RL Biochem. Biophys. Res. Commun. 158:906-912(1989).  
 RN [8]  
 RP SEQUENCE OF 289-364 FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Placenta;  
 RX MEDLINE=89345111; PubMed=2569710;  
 RA Fukuchi K., Martin G.M., Deeb S.S.;  
 RT "Sequence of the protease inhibitor domain of the A4 amyloid protein  
 RT precursor of Mus domesticus.";  
 RL Nucleic Acids Res. 17:5396-5396(1989).  
 RN [9]  
 RP SEQUENCE OF 656-737 FROM N.A.  
 RC STRAIN=129/Sv;  
 RA Wragg M.A., Busfield F., Duff K., Korenblat K., Capecchi M.,  
 RA Loring J.F., Goate A.M.;  
 RT "Introduction of six mutations into the mouse genome using 'Hit and  
 RT Run' gene-targeting: introduction of familial Alzheimer's disease  
 RT mutations into the mouse amyloid precursor protein gene and  
 RT humanization of the A-beta fragment.";  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.  
 RN [10]  
 RP TISSUE SPECIFICITY OF ALTERNATIVE SPLICED FORMS.  
 RX MEDLINE=93287808; PubMed=8510506;  
 RA Sola C., Mengod G., Ghetti B., Palacios J.M., Triarhou L.C.;  
 RT "Regional distribution of the alternatively spliced isoforms of beta  
 RT APP RNA transcript in the brain of normal, heterozygous and  
 RT homozygous weaver mutant mice as revealed by in situ hybridization  
 RT histochemistry.";  
 RL Brain Res. Mol. Brain Res. 17:340-346(1993).  
 RN [11]  
 RP INTERACTION WITH KNS2.  
 RX MEDLINE=21010507; PubMed=11144355;  
 RA Kamal A., Stokin G.B., Yang Z., Xia C.-H., Goldstein L.S.;  
 RT "Axonal transport of amyloid precursor protein is mediated by direct  
 RT binding to the kinesin light chain subunit of kinesin-I.";  
 RL Neuron 28:449-459(2000).  
 RN [12]  
 RP C-TERMINAL PROTEIN-PROTEIN INTERACTIONS, AND MUTAGENESIS OF TYR-728;  
 RP THR-743; TYR-757; ASN-759 AND TYR-762.  
 RX MEDLINE=21408156; PubMed=11517249;  
 RA Matsuda S., Yasukawa T., Homma Y., Ito Y., Niikura T., Hiraki T.,  
 RA Hirai S., Ohno S., Kita Y., Kawasumi M., Kouyama K., Yamamoto T.,

RA Kyriakis J.M., Nishimoto I.;  
 RT "C-jun N-terminal kinase (JNK)-interacting protein-1b/islet-brain-1  
 RT scaffolds Alzheimer's amyloid precursor protein with JNK.";  
 RL J. Neurosci. 21:6597-6607(2001).  
 RN [13]  
 RP INTERACTION WITH MAPK8IP1, AND PHOSPHORYLATION.  
 RX MEDLINE=22028091; PubMed=11912189;  
 RA Taru H., Iijima K.-I., Hase M., Kirino Y., Yagi Y., Suzuki T.;  
 RT "Interaction of Alzheimer's beta-amyloid precursor family proteins  
 RT with scaffold proteins of the JNK signaling cascade.";  
 RL J. Biol. Chem. 277:20070-20078(2002).  
 RN [14]  
 RP INTERACTION OF CTF PEPTIDES WITH NUMB.  
 RX MEDLINE=22008109; PubMed=12011466;  
 RA Roncarati R., Sestan N., Scheinfeld M.H., Berechid B.E., Lopez P.A.,  
 RA Meucci O., McGlade J.C., Rakic P., D'Adamio L.;  
 RT "The gamma-secretase-generated intracellular domain of beta-amyloid  
 RT precursor protein binds Numb and inhibits Notch signaling.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:7102-7107(2002).  
 RN [15]  
 RP GAMMA-SECRETASE PROCESSING, AND INTERACTION WITH APBB1.  
 RX MEDLINE=21437805; PubMed=11553691;  
 RA Cupers P., Orlans I., Craessaerts K., Annaert W., De Strooper B.;  
 RT "The amyloid precursor protein (APP)-cytoplasmic fragment generated by  
 RT gamma-secretase is rapidly degraded but distributes partially in a  
 RT nuclear fraction of neurones in culture.";  
 RL J. Neurochem. 78:1168-1178(2001).  
 CC -!- FUNCTION: Functions as a cell surface receptor and performs  
 CC physiological functions on the surface of neurons relevant to  
 CC neurite growth, neuronal adhesion and axonogenesis. Involved in  
 CC cell mobility and transcription regulation through protein-protein  
 CC interactions. Can promote transcription activation through binding  
 CC to APBB1/Tip60 and inhibit Notch signaling through interaction  
 CC with Numb. Couples to apoptosis-inducing pathways such as those  
 CC mediated by G(0) and JIP. Inhibits G(0) alpha ATPase activity (By  
 CC similarity). Acts as a kinesin I membrane receptor, mediating the  
 CC axonal transport of beta-secretase and presenilin 1. May be  
 CC involved in copper homeostasis/oxidative stress through copper ion  
 CC reduction. Can regulate neurite outgrowth through binding to  
 CC components of the extracellular matrix such as heparin and  
 CC collagen I and IV (By similarity). The splice isoforms that  
 CC contain the BPTI domain possess protease inhibitor activity (By  
 CC similarity).  
 CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators  
 CC with metal-reducing activity. Bind transient metals such as  
 CC copper, zinc and iron. Rat and mouse beta-amyloid peptides bind  
 CC only weakly transient metals and have little reducing activity due  
 CC to substitutions of transient metal chelating residues. Beta-APP42  
 CC may activate mononuclear phagocytes in the brain and elicit  
 CC inflammatory responses. Promotes both tau aggregation and TPK II-  
 CC mediated phosphorylation (By similarity).  
 CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved  
 CC peptides, including C31, are potent enhancers of neuronal  
 CC apoptosis.  
 CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several  
 CC cytoplasmic proteins, including APBB family members, the APBA  
 CC family, MAPK8IP1, SHC1, Numb and Dab1. Binding to Dab1 inhibits

CC its serine phosphorylation. Also interacts with GPCR-like protein  
 CC BPP, FPRL1, APPBP1, IBL, KNS2 (via its TPR domains), APPBP2 (via  
 CC BaSS) and DDB1 (By similarity). In vitro, it binds MAPT via the  
 CC MT-binding domains (By similarity). Associates with microtubules  
 CC in the presence of ATP and in a kinesin-dependent manner (By  
 CC similarity). Interacts, through a C-terminal domain, with GNAO1  
 CC (By similarity). Amyloid beta-42 binds CHRNA7 in hippocampal  
 CC neurons (By similarity). Beta-amyloid associates with HADH2 (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface  
 CC protein that rapidly becomes internalized via clathrin-coated  
 CC pits. During maturation, the immature APP (N-glycosylated in the  
 CC endoplasmic reticulum) moves to the Golgi complex where complete

Query Match 100.0%; Score 40; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 687 KLVFFAED 694

#### RESULT 11

##### A4\_PIG

ID A4\_PIG STANDARD; PRT; 770 AA.  
 AC P79307; Q29023; Q9TUI0;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease  
 DE amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);  
 DE Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-  
 DE APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40);  
 DE Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)  
 DE (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-  
 DE secretase C-terminal fragment 50); C31].  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kimura A., Takahashi T.;  
 RT "Amyloid precursor protein 770.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-136 FROM N.A.  
 RC TISSUE=Small intestine;  
 RA Winteroe A.K., Fredholm M.;  
 RT "Evaluation and characterization of a porcine small intestine cDNA  
 RT library.";  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE OF 667-723 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=92017079; PubMed=1656157;

RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;  
 RT "Conservation of the sequence of the Alzheimer's disease amyloid  
 RT peptide in dog, polar bear and five other mammals by cross-species  
 RT polymerase chain reaction analysis.";  
 RL Brain Res. Mol. Brain Res. 10:299-305(1991).

CC -!- FUNCTION: Functions as a cell surface receptor and performs  
 CC physiological functions on the surface of neurons relevant to  
 CC neurite growth, neuronal adhesion and axonogenesis. Involved in  
 CC cell mobility and transcription regulation through protein-protein  
 CC interactions (By similarity). Can promote transcription activation  
 CC through binding to APBB1/Tip60 and inhibit Notch signaling through  
 CC interaction with Numb (By similarity). Couples to apoptosis-  
 CC inducing pathways such as those mediated by G(O) and JIP (By  
 CC similarity). Inhibits G(O) alpha ATPase activity (By similarity).  
 CC Acts as a kinesin I membrane receptor, mediating the axonal  
 CC transport of beta-secretase and presenilin 1 (By similarity). May  
 CC be involved in copper homeostasis/oxidative stress through copper  
 CC ion reduction (By similarity). In vitro, copper-metallated APP  
 CC induces neuronal death directly or is potentiated through Cu(II)-  
 CC mediated low-density lipoprotein oxidation (By similarity). Can  
 CC regulate neurite outgrowth through binding to components of the  
 CC extracellular matrix such as heparin and collagen I and IV (By  
 CC similarity).

CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators  
 CC with metal-reducing activity. Bind transient metals such as  
 CC copper, zinc and iron (By similarity).

CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved  
 CC peptides, including C31, are potent enhancers of neuronal  
 CC apoptosis (By similarity).

CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several  
 CC cytoplasmic proteins, including APBB family members, the APBA  
 CC family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding  
 CC to Dab1 inhibits its serine phosphorylation (By similarity). Also  
 CC interacts with GPCR-like protein BPP, FPRL1, APPBP1, IBL, KNS2  
 CC (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.  
 CC In vitro, it binds MAPT via the MT-binding domains (By  
 CC similarity). Associates with microtubules in the presence of ATP  
 CC and in a kinesin-dependent manner (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface  
 CC protein that rapidly becomes internalized via clathrin-coated  
 CC pits. During maturation, the immature APP (N-glycosylated in the  
 CC endoplasmic reticulum) moves to the Golgi complex where complete  
 CC maturation occurs (O-glycosylated and sulfated). After alpha-  
 CC secretase cleavage, soluble APP is released into the extracellular  
 CC space and the C-terminal is internalized to endosomes and  
 CC lysosomes. Some APP accumulates in secretory transport vesicles  
 CC leaving the late Golgi compartment and returns to the cell  
 CC surface. Gamma-CTF(59) peptide is located to both the cytoplasm  
 CC and nuclei of neurons (By similarity).

CC -!- DOMAIN: The basolateral sorting signal (BaSS) is required for  
 CC sorting of membrane proteins to the basolateral surface of  
 CC epithelial cells (By similarity).

CC -!- DOMAIN: The NPXY sequence motif found in many tyrosine-  
 CC phosphorylated proteins is required for the specific binding of  
 CC the PID domain. However additional amino acids either N- or C-  
 CC terminal to the NPXY motif are often required for complete  
 CC interaction. The PID domain-containing proteins which bind APP

CC require the YENPTY motif for full interaction. These interactions  
 CC are independent of phosphorylation on the terminal tyrosine  
 CC residue. The NPXY site is also involved in clathrin-mediated  
 CC endocytosis (By similarity).  
 CC -!- PTM: Proteolytically processed under normal cellular conditions.  
 CC Cleavage by alpha-secretase or alternatively by beta-secretase  
 CC leads to generation and extracellular release of soluble APP  
 CC peptides, S-APP-alpha and S-APP-beta, respectively, and the  
 CC retention of corresponding membrane-anchored C-terminal fragments,  
 CC C83 and C99. Subsequent processing of C83 by gamma-secretase  
 CC yields P3 peptides. This is the major secretory pathway and is  
 CC nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated  
 CC gamma-secretase processing of C99 releases the amyloid beta  
 CC proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),  
 CC major components of amyloid plaques, and the cytotoxic C-terminal  
 CC fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By  
 CC similarity).  
 CC -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis  
 CC (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9  
 CC results in the production of the neurotoxic C31 peptide and the  
 CC increased production of beta-amyloid peptides (By similarity).  
 CC -!- PTM: N- and O-glycosylated (By similarity).  
 CC -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and  
 CC serine residues is neuron-specific. Phosphorylation can affect APP  
 CC processing, neuronal differentiation and interaction with other  
 CC proteins (By similarity).  
 CC -!- PTM: Extracellular binding and reduction of copper, results in a  
 CC corresponding oxidation of Cys-144 and Cys-158, and the formation  
 CC of a disulfide bond (By similarity).  
 CC -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and  
 CC zinc, can induce histidine-bridging between beta-amyloid molecules  
 CC resulting in beta-amyloid-metal aggregates (By similarity).  
 CC Extracellular zinc-binding increases binding of heparin to APP and  
 CC inhibits collagen-binding (By similarity).  
 CC -!- SIMILARITY: Belongs to the APP family.  
 CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AB032550; BAA84580.1; -.  
 DR EMBL; Z84022; CAB06313.1; -.  
 DR EMBL; X56127; CAA39592.1; -.  
 DR HSSP; P05067; 1AAP.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00006; A4\_EXTRA; 1.

DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 KW Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;  
 KW Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;  
 KW Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;  
 KW Amyloid.

FT	SIGNAL	1	17	BY SIMILARITY.
FT	CHAIN	18	770	AMYLOID BETA A4 PROTEIN.
FT	CHAIN	18	687	SOLUBLE APP-ALPHA (POTENTIAL).
FT	CHAIN	18	671	SOLUBLE APP-BETA (POTENTIAL).
FT	CHAIN	672	770	C99 (BY SIMILARITY).
FT	CHAIN	672	713	BETA-AMYLOID PROTEIN 42 (BY SIMILARITY).
FT	CHAIN	672	711	BETA-AMYLOID PROTEIN 40 (BY SIMILARITY).
FT	CHAIN	688	770	C83 (BY SIMILARITY).
FT	CHAIN	688	713	P3(42) (BY SIMILARITY).
FT	CHAIN	688	711	P3(40) (BY SIMILARITY).
FT	CHAIN	712	770	GAMMA-CTF(59).
FT	CHAIN	714	770	GAMMA-CTF(57).
FT	CHAIN	721	770	GAMMA-CTF(50) (BY SIMILARITY).
FT	CHAIN	740	770	C31 (DURING APOPTOSIS) (BY SIMILARITY).
FT	DOMAIN	18	699	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	700	723	POTENTIAL.
FT	DOMAIN	724	770	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	96	110	HEPARIN-BINDING (BY SIMILARITY).
FT	DOMAIN	135	155	COPPER-BINDING (BY SIMILARITY).
FT	DOMAIN	181	188	ZINC-BINDING (BY SIMILARITY).
FT	DOMAIN	291	341	BPTI/KUNITZ INHIBITOR.
FT	DOMAIN	391	423	HEPARIN-BINDING (BY SIMILARITY).
FT	DOMAIN	491	522	HEPARIN-BINDING (BY SIMILARITY).
FT	DOMAIN	523	540	COLLAGEN-BINDING (BY SIMILARITY).
FT	DOMAIN	732	751	INTERACTION WITH G(O)-ALPHA (BY SIMILARITY).
FT				
FT	DOMAIN	230	260	ASP/GLU-RICH (ACIDIC).
FT	DOMAIN	274	280	POLY-THR.
FT	SITE	144	144	REQUIRED FOR COPPER(II) REDUCTION (BY SIMILARITY).
FT				
FT	ACT_SITE	301	302	REACTIVE BOND (BY SIMILARITY).
FT	SITE	671	672	CLEAVAGE (BY BETA-SECRETASE) (BY SIMILARITY).
FT				
FT	SITE	672	673	CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT	SITE	687	688	CLEAVAGE (BY ALPHA-SECRETASE) (BY SIMILARITY).
FT				
FT	SITE	704	704	IMPLICATED IN FREE RADICAL PROPAGATION (BY SIMILARITY).
FT				
FT	SITE	706	706	INVOLVED IN OXIDATIVE REACTIONS (BY SIMILARITY).
FT				
FT	SITE	711	712	CLEAVAGE (BY GAMMA-SECRETASE; SITE 1) (BY SIMILARITY).
FT				
FT	SITE	713	714	CLEAVAGE (BY GAMMA-SECRETASE; SITE 2) (BY SIMILARITY).
FT				
FT	SITE	720	721	CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)

Query Match 100.0%; Score 40; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 0.38;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 687 KLVFFAED 694

RESULT 12

A4\_RAT

ID A4\_RAT STANDARD; PRT; 770 AA.  
 AC P08592;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Amyloid beta A4 protein precursor (APP) (Alzheimer's disease amyloid  
 DE protein homolog) (Amyloidogenic glycoprotein) (AG) [Contains: Soluble  
 DE APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99; Beta-  
 DE amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40);  
 DE C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal  
 DE fragment 59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57);  
 DE Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50); C31].  
 GN APP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM APP695).  
 RC TISSUE=Brain;  
 RX MEDLINE=88312583; PubMed=2900758;  
 RA Shivers B.D., Hilbich C., Multhaup G., Salbaum J.M., Beyreuther K.,  
 RA Seeburg P.H.;  
 RT "Alzheimer's disease amyloidogenic glycoprotein: expression pattern  
 RT in rat brain suggests a role in cell contact.";  
 RL EMBO J. 7:1365-1370(1988).  
 RN [2]  
 RP SEQUENCE OF 289-364 FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=89183625; PubMed=2648331;  
 RA Kang J., Mueller-Hill B.;  
 RT "The sequence of the two extra exons in rat preA4.";  
 RL Nucleic Acids Res. 17:2130-2130(1989).  
 RN [3]  
 RP SEQUENCE OF 720-730, AND MASS SPECTROMETRY.  
 RX MEDLINE=21443797; PubMed=11483588;  
 RA Gu Y., Misonou H., Sato T., Dohmae N., Takio K., Ihara Y.;  
 RT "Distinct intramembrane cleavage of the beta-amyloid precursor protein  
 RT family resembling gamma-secretase-like cleavage of Notch.";  
 RL J. Biol. Chem. 276:35235-35238(2001).  
 RN [4]  
 RP ALTERNATIVE SPLICING.  
 RX MEDLINE=96187032; PubMed=8624099;  
 RA Sandbrink R., Masters C.L., Beyreuther K.;  
 RT "APP gene family. Alternative splicing generates functionally related  
 RT isoforms.";  
 RL Ann. N.Y. Acad. Sci. 777:281-287(1996).  
 RN [5]

RP TISSUE SPECIFICITY OF APPICAN.  
RX MEDLINE=95263526; PubMed=7744833;  
RA Shioi J., Pangalos M.N., Ripellino J.A., Vassilacopoulou D.,  
RA Mytilineou C., Margolis R.U., Robakis N.K.;  
RT "The Alzheimer amyloid precursor proteoglycan (appican) is present in  
RT brain and is produced by astrocytes but not by neurons in primary  
RT neural cultures.";  
RL J. Biol. Chem. 270:11839-11844(1995).  
RN [6]

RP TISSUE SPECIFICITY OF ISOFORMS.  
RX MEDLINE=97150061; PubMed=8996834;  
RA Sandbrink R., Monning U., Masters C.L., Beyreuther K.;  
RT "Expression of the APP gene family in brain cells, brain development  
RT and aging.";  
RL Gerontology 43:119-131(1997).  
RN [7]

RP INTERACTION WITH DDB1, AND MUTAGENESIS OF TYR-757; ASN-759 AND  
RP TYR-762.  
RX MEDLINE=99127916; PubMed=9930726;  
RA Watanabe T., Sukegawa J., Tomita S., Iijima K.-I., Oguchi S.,  
RA Suzuki T., Nairn A.C., Greengard P.;  
RT "A 127-kDa protein (UV-DDB) binds to the cytoplasmic domain of the  
RT Alzheimer's amyloid precursor protein.";  
RL J. Neurochem. 72:549-556(1999).  
RN [8]

RP INTERACTION WITH GNAO1, AND MUTAGENESIS OF 732-HIS-HIS-733.  
RX MEDLINE=99162676; PubMed=10024358;  
RA Brouillet E., Trembleau A., Galanaud D., Volovitch M., Bouillot C.,  
RA Valenza C., Prochiantz A., Allinquant B.;  
RT "The amyloid precursor protein interacts with Go heterotrimeric  
RT protein within a cell compartment specialized in signal  
RT transduction.";  
RL J. Neurosci. 19:1717-1727(1999).  
RN [9]

RP CHARACTERISTICS OF APPICAN, AND MUTAGENESIS OF SER-656.  
RX MEDLINE=95256193; PubMed=7737970;  
RA Pangalos M.N., Efthimiopoulos S., Shioi J., Robakis N.K.;  
RT "The chondroitin sulfate attachment site of appican is formed by  
RT splicing out exon 15 of the amyloid precursor gene.";  
RL J. Biol. Chem. 270:10388-10391(1995).  
RN [10]

RP BETA-AMYLOID METAL-BINDING.  
RX MEDLINE=99316162; PubMed=10386999;  
RA Huang X., Atwood C.S., Hartshorn M.A., Multhaup G., Goldstein L.E.,  
RA Scarpa R.C., Cuajungco M.P., Gray D.N., Lim J., Moir R.D., Tanzi R.E.,  
RA Bush A.I.;  
RT "The A beta peptide of Alzheimer's disease directly produces hydrogen  
RT peroxide through metal ion reduction.";  
RL Biochemistry 38:7609-7616(1999).  
RN [11]

RP BETA-AMYLOID ZINC BINDING.  
RX MEDLINE=99343552; PubMed=10413512;  
RA Liu S.T., Howlett G., Barrow C.J.;  
RT "Histidine-13 is a crucial residue in the zinc ion-induced aggregation  
RT of the A beta peptide of Alzheimer's disease.";  
RL Biochemistry 38:9373-9378(1999).  
RN [12]



RP IMPORTANCE OF GLY-704 IN FREE RADICAL PROPAGATION, AND MUTAGENESIS OF  
 RP GLY-704.  
 RX MEDLINE=21956095; PubMed=11959460;  
 RA Kanski J., Varadarajan S., Aksenova M., Butterfield D.A.;  
 RT "Role of glycine-33 and methionine-35 in Alzheimer's amyloid beta-  
 RT peptide 1-42-associated oxidative stress and neurotoxicity.";  
 RL Biochim. Biophys. Acta 1586:190-198(2001).  
 RN [13]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=97239592; PubMed=9085254;  
 RA Oishi M., Nairn A.C., Czernik A.J., Lim G.S., Isohara T., Gandy S.E.,  
 RA Greengard P., Suzuki T.;  
 RT "The cytoplasmic domain of Alzheimer's amyloid precursor protein is  
 RT phosphorylated at Thr654, Ser655, and Thr668 in adult rat brain and  
 RT cultured cells.";  
 RL Mol. Med. 3:111-123(1997).  
 RN [14]  
 RP PHOSPHORYLATION ON SER-730.  
 RX MEDLINE=99262094; PubMed=10329382;  
 RA Isohara T., Horiuchi A., Watanabe T., Ando K., Czernik A.J., Uno I.,  
 RA Greengard P., Nairn A.C., Suzuki T.;  
 RT "Phosphorylation of the cytoplasmic domain of Alzheimer's beta-amyloid  
 RT precursor protein at Ser655 by a novel protein kinase.";  
 RL Biochem. Biophys. Res. Commun. 258:300-305(1999).  
 RN [15]  
 RP PHOSPHORYLATION, INDUCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF  
 RP THR-743.  
 RX MEDLINE=99274744; PubMed=10341243;  
 RA Ando K., Oishi M., Takeda S., Iijima K.-I., Isohara T., Nairn A.C.,  
 RA Kirino Y., Greengard P., Suzuki T.;  
 RT "Role of phosphorylation of Alzheimer's amyloid precursor protein  
 RT during neuronal differentiation.";  
 RL J. Neurosci. 19:4421-4427(1999).  
 RN [16]  
 RP PHOSPHORYLATION ON THR-743.  
 RX MEDLINE=20396183; PubMed=10936190;  
 RA Iijima K.-I., Ando K., Takeda S., Satoh Y., Seki T., Itohara S.,  
 RA Greengard P., Kirino Y., Nairn A.C., Suzuki T.;  
 RT "Neuron-specific phosphorylation of Alzheimer's beta-amyloid precursor  
 RT protein by cyclin-dependent kinase 5.";  
 RL J. Neurochem. 75:1085-1091(2000).  
 RN [17]  
 RP CARBOHYDRATE STRUCTURE OF APPICAN.  
 RX MEDLINE=21463085; PubMed=11479316;  
 RA Tsuchida K., Shioi J., Yamada S., Boghosian G., Wu A., Cai H.,  
 RA Sugahara K., Robakis N.K.;  
 RT "Appican, the proteoglycan form of the amyloid precursor protein,  
 RT contains chondroitin sulfate E in the repeating disaccharide region  
 RT and 4-O-sulfated galactose in the linkage region.";  
 RL J. Biol. Chem. 276:37155-37160(2001).  
 CC -!- FUNCTION: Functions as a cell surface receptor and performs  
 CC physiological functions on the surface of neurons relevant to  
 CC neurite growth, neuronal adhesion and axonogenesis. Involved in  
 CC cell mobility and transcription regulation through protein-protein  
 CC interactions (By similarity). Can promote transcription activation  
 CC through binding to APBB1/Tip60 and inhibit Notch signaling through  
 CC interaction with Numb (By similarity). Couples to apoptosis-

CC inducing pathways such as those mediated by G(0) and JIP. Inhibits  
 CC G(0) alpha ATPase activity. Acts as a kinesin I membrane receptor,  
 CC mediating the axonal transport of beta-secretase and presenilin 1  
 CC (By similarity). May be involved in copper homeostasis/oxidative  
 CC stress through copper ion reduction. Can regulate neurite  
 CC outgrowth through binding to components of the extracellular  
 CC matrix such as heparin and collagen I and IV (By similarity). The  
 CC splice isoforms that contain the BPTI domain possess protease  
 CC inhibitor activity (By similarity).  
 CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators  
 CC with metal-reducing activity. Bind transient metals such as  
 CC copper, zinc and iron. Rat and mouse beta-amyloid peptides bind  
 CC only weakly transient metals and have little reducing activity due  
 CC to substitutions of transient metal chelating residues. Beta-APP42  
 CC may activate mononuclear phagocytes in the brain and elicit  
 CC inflammatory responses. Promotes both tau aggregation and TPK II-  
 CC mediated phosphorylation (By similarity).  
 CC -!- FUNCTION: Appicans elicit adhesion of neural cells to the  
 CC extracellular matrix and may regulate neurite outgrowth in the  
 CC brain.  
 CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved  
 CC peptides, including C31, are potent enhancers of neuronal  
 CC apoptosis (By similarity).  
 CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several  
 CC cytoplasmic proteins, including APBB family members, the APBA  
 CC family, MAPK8IP1, SHC1 and Numb and Dab1 (By similarity). Binding  
 CC to Dab1 inhibits its serine phosphorylation (By similarity). Also  
 CC interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2  
 CC (via its TPR domains), APPBP2 (via BaSS) (By similarity) and DDB1.  
 CC In vitro, it binds MAPT via the MT-binding domains (By  
 CC similarity). Associates with microtubules in the presence of ATP  
 CC and in a kinesin-dependent manner (By similarity). Interacts,  
 CC through a C-terminal domain, with GNAO1. Amyloid beta-42 binds  
 CC CHRNA7 in hippocampal neurons (By similarity). Beta-amyloid  
 CC associates with HADH2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface  
 CC protein that rapidly becomes internalized via clathrin-coated  
 CC pits. During maturation, the immature APP (N-glycosylated in the

Query Match 100.0%; Score 40; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 687 KLVFFAED 694

#### RESULT 13

##### A4\_TETFL

ID A4\_TETFL STANDARD; PRT; 780 AA.  
 AC 073683;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alzheimer's disease amyloid A4 protein homolog precursor [Contains:  
 DE Beta-amyloid protein (Beta-APP) (A-beta)].

GN APP.  
 OS Tetraodon fluviatilis (Puffer fish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontoidea; Tetraodontidae; Tetraodon.  
 OX NCBI\_TaxID=47145;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98252138; PubMed=9599080;  
 RA Villard L., Tassone F., Crnogorac-Jurcevic T., Clancy K., Gardiner K.;  
 RT "Analysis of pufferfish homologues of the AT-rich human APP gene."  
 RL Gene 210:17-24(1998).  
 CC -!- FUNCTION: Functional neuronal receptor which couples to  
 CC intracellular signaling pathway through the GTP-binding protein  
 CC G(O) (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Belongs to the APP family.  
 CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF018165; AAC41275.1; -.  
 DR HSSP; P05067; 1HZ3.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR Pfam; PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; FALSE\_NEG.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 KW Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;  
 KW Serine protease inhibitor.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 780 ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN  
 FT HOMOLOG.  
 FT CHAIN 682 724 BETA-AMYLOID PROTEIN (POTENTIAL).  
 FT DOMAIN 19 711 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 712 732 POTENTIAL.  
 FT DOMAIN 733 780 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 323 382 BPTI/KUNITZ INHIBITOR.  
 FT SITE 769 772 CLATHRIN-BINDING (BY SIMILARITY).

FT DISULFID 327 378 BY SIMILARITY.  
 FT DISULFID 336 361 BY SIMILARITY.  
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 780 AA; 88238 MW; 60071BE94520191D CRC64;

Query Match 100.0%; Score 40; DB 1; Length 780;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||||  
 Db 697 KLVFFAED 704

# RESULT 14

## A4\_FUGRU

ID A4\_FUGRU STANDARD; PRT; 737 AA.  
 AC O93279;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alzheimer's disease amyloid A4 protein homolog precursor [Contains:  
 DE Beta-amyloid protein (Beta-APP) (A-beta)].  
 GN APP.  
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetradontoidea; Tetraodontidae; Takifugu.  
 OX NCBI\_TaxID=31033;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98252138; PubMed=9599080;  
 RA Villard L., Tassone F., Crnogorac-Jurcevic T., Clancy K., Gardiner K.;  
 RT "Analysis of pufferfish homologues of the AT-rich human APP gene."  
 RL Gene 210:17-24(1998).  
 CC -!- FUNCTION: Functional neuronal receptor which couples to  
 CC intracellular signaling pathway through the GTP-binding protein  
 CC G(O) (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Belongs to the APP family.  
 CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF090120; AAD13392.1; -.  
 DR HSSP; P05067; 1HZ3.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR InterPro; IPR002223; Kunitz\_BPTI.

DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR Pfam; PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; FALSE\_NEG.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 KW Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;  
 KW Serine protease inhibitor.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 737 ALZHEIMER'S DISEASE AMYLOID A4  
 FT PROTEIN HOMOLOG.  
 FT CHAIN 639 681 BETA-AMYLOID PROTEIN (POTENTIAL).  
 FT DOMAIN 19 668 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 669 689 POTENTIAL.  
 FT DOMAIN 690 737 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 286 344 BPTI/KUNITZ INHIBITOR.  
 FT SITE 726 729 CLATHRIN-BINDING (BY SIMILARITY).  
 FT ACT\_SITE 300 301 REACTIVE BOND.  
 FT DISULFID 290 340 BY SIMILARITY.  
 FT DISULFID 299 323 BY SIMILARITY.  
 FT DISULFID 315 336 BY SIMILARITY.  
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 737 AA; 82856 MW; 6FAD01E2E3B2B7E2 CRC64;

Query Match 92.5%; Score 37; DB 1; Length 737;  
 Best Local Similarity 87.5%; Pred. No. 1.7;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||:|  
 Db 654 KLVFFADD 661

# RESULT 15

Y189\_RICPR

ID Y189\_RICPR STANDARD; PRT; 321 AA.  
 AC Q9ZDX5;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein RP189.  
 GN RP189.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
 OC Rickettsiaceae; Rickettsieae; Rickettsia.  
 OX NCBI\_TaxID=782;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Madrid E;  
 RX MEDLINE=99039499; PubMed=9823893;  
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,

RA Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RT mitochondria.";  
 RL Nature 396:133-140(1998).  
 CC -!- SIMILARITY: SOME, TO A.AEOLICUS AQ\_1104.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ235270; CAA14655.1; -.  
 DR PIR; H71729; H71729.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 321 AA; 36653 MW; 3E5F47D104DD8A73 CRC64;

Query Match 85.0%; Score 34; DB 1; Length 321;  
 Best Local Similarity 75.0%; Pred. No. 3.7;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||:||||  
 Db 178 KLIFFAHD 185

# RESULT 16

RL7A\_ARCFU

ID RL7A\_ARCFU STANDARD; PRT; 119 AA.

AC 029494;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 50S ribosomal protein L7Ae.

GN RPL7AE OR AF0764.

OS Archaeoglobus fulgidus.

OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;

OC Archaeoglobaceae; Archaeoglobus.

OX NCBI\_TaxID=2234;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;

RX MEDLINE=98049343; PubMed=9389475;

RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,

RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,

RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,

RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,

RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,

RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,

RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,

RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,

RA Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-  
RT reducing archaeon *Archaeoglobus fulgidus*.";  
RL Nature 390:364-370(1997).  
CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE001051; AAB90466.1; -.  
DR PIR; D69345; D69345.  
DR HSSP; P55769; 1E7K.  
DR TIGR; AF0764; -.  
DR HAMAP; MF\_00326; -; 1.  
DR InterPro; IPR004038; Ribosomal\_L7A.  
DR InterPro; IPR004037; Ribosomal\_L7Ae.  
DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
DR PRINTS; PR00881; L7ARS6FAMILY.  
DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 119 AA; 13193 MW; 9065A83C82207501 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 119;  
Best Local Similarity 75.0%; Pred. No. 6.7;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | : | | |  
Db 45 KLVYIAED 52

# RESULT 17

## UPK\_CORST

ID UPK\_CORST STANDARD; PRT; 281 AA.  
AC Q9FB58;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Putative undecaprenol kinase (EC 2.7.1.66) (Bacitracin resistance  
DE protein).  
GN UPK OR BACA.  
OS *Corynebacterium striatum*.  
OG Plasmid pTP10.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC *Corynebacterineae*; *Corynebacteriaceae*; *Corynebacterium*.  
OX NCBI\_TaxID=43770;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M82B;  
RX MEDLINE=20194806; PubMed=10732668;  
RA Tauch A., Krieft S., Kalinowski J., Puehler A.;  
RT "The 51,409-bp R-plasmid pTP10 from the multiresistant clinical  
RT isolate *Corynebacterium striatum* M82B is composed of DNA segments

RT initially identified in soil bacteria and in plant, animal, and human  
RT pathogens.";  
RL Mol. Gen. Genet. 263:1-11(2000).  
CC -!- FUNCTION: Probably phosphorylates undecaprenol to undecaprenyl  
CC phosphate. Confers resistance to bacitracin (By similarity).  
CC -!- CATALYTIC ACTIVITY: ATP + undecaprenol = ADP + undecaprenyl  
CC phosphate.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -!- MISCELLANEOUS: Bacitracin is thought to be involved in inhibition  
CC of peptidoglycan synthesis by sequestering undecaprenyl  
CC diphosphate reducing the pool of lipid carrier available.  
CC -!- SIMILARITY: Belongs to the upk family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF024666; AAG03365.1; -.  
DR HAMAP; MF\_01006; -; 1.  
DR InterPro; IPR003824; BacA.  
DR Pfam; PF02673; BacA; 1.  
DR TIGRFAMs; TIGR00753; undec\_kin\_bacA; 1.  
KW Transferase; Kinase; Antibiotic resistance; Transmembrane; Plasmid.  
FT TRANSMEM 13 32 POTENTIAL.  
FT TRANSMEM 52 71 POTENTIAL.  
FT TRANSMEM 98 115 POTENTIAL.  
FT TRANSMEM 125 142 POTENTIAL.  
FT TRANSMEM 198 220 POTENTIAL.  
FT TRANSMEM 230 252 POTENTIAL.  
FT TRANSMEM 259 278 POTENTIAL.  
SQ SEQUENCE 281 AA; 30512 MW; F82700075DB27EA9 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 281;  
Best Local Similarity 85.7%; Pred. No. 15;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LVFFAED 8  
| | | | | : |  
Db 67 LVFFAKD 73

# RESULT 18

## LICR\_BACSU

ID LICR\_BACSU STANDARD; PRT; 641 AA.  
AC P46321;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Probable licABCH operon regulator.  
GN LICR OR CELR OR BSU38600.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;



RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=97144536; PubMed=8990303;  
 RA Tobisch S., Glaser P., Krueger S., Hecker M.;  
 RT "Identification and characterization of a new beta-glucoside  
 RT utilization system in *Bacillus subtilis*."  
 RL J. Bacteriol. 179:496-506(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / BGSC1A1;  
 RX MEDLINE=97124196; PubMed=8969509;  
 RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,  
 RA Miwa Y., Fujita Y.;  
 RT "Sequencing of a 65 kb region of the *Bacillus subtilis* genome  
 RT containing the *lic* and *cel* loci, and creation of a 177 kb contig  
 RT covering the *gnt-sacXY* region."  
 RL Microbiology 142:3113-3123(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*."  
 RL Nature 390:249-256(1997).  
 RN [4]  
 RP MUTAGENESIS OF HIS-219; HIS-278; HIS-333; HIS-392 AND HIS-559.  
 RX MEDLINE=99369874; PubMed=10438772;  
 RA Tobisch S., Stuelke J., Hecker M.;

```

RT "Regulation of the lic operon of Bacillus subtilis and
RT characterization of potential phosphorylation sites of the LicR
RT regulator protein by site-directed mutagenesis.";
RL J. Bacteriol. 181:4995-5003(1999).
CC -!- FUNCTION: Positive regulator of the licABCH operon.
CC -!- MISCELLANEOUS: Might be activated by phosphorylation of certain
CC histidyl residues by HPr.
CC -!- SIMILARITY: Contains 1 PTS EIIA domain.
CC -!- SIMILARITY: Belongs to the transcriptional antiterminator bglG
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49992; CAA90284.1; -.
DR EMBL; D83026; BAA11742.1; -.
DR EMBL; Z99123; CAB15886.1; -.
DR PIR; H69651; H69651.
DR SubtiList; Bgl1346; licR.
DR InterPro; IPR001550; BglG_antitermin.
DR InterPro; IPR002178; PTS_EIIA_2.
DR Pfam; PF00874; PRD; 2.
DR Pfam; PF00359; PTS_EIIA_2; 1.
DR ProDom; PD001689; PTS_EIIA_2; 1.
DR PROSITE; PS00654; ANTITERMINATORS_BGLG; FALSE_NEG.
KW Transcription regulation; Activator; RNA-binding; Phosphorylation;
KW Complete proteome.
FT DOMAIN 499 638 EIIA DOMAIN.
FT MOD_RES 219 219 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 278 278 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 333 333 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 392 392 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 559 559 PHOSPHORYLATION (POTENTIAL).
FT MUTAGEN 219 219 H->D: LOSS OF ACTIVITY.
FT MUTAGEN 219 219 H->A: RESIDUAL ACTIVITY.
FT MUTAGEN 278 278 H->E: LOSS OF ACTIVITY.
FT MUTAGEN 278 278 H->A: LOSS OF ACTIVITY.
FT MUTAGEN 333 333 H->A: LOSS OF ACTIVITY.
FT MUTAGEN 392 392 H->E: LOSS OF ACTIVITY.
FT MUTAGEN 392 392 H->I: RESIDUAL ACTIVITY.
FT MUTAGEN 559 559 H->G: INCREASE IN ACTIVITY.
SQ SEQUENCE 641 AA; 73315 MW; FCEF83BFC72A0168 CRC64;

```

```

Query Match 77.5%; Score 31; DB 1; Length 641;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 KLVFFAED 8
   :|||| ||
Db 504 ELVFFQED 511

```

RESULT 19

MOR2\_SCHPO

ID MOR2\_SCHPO STANDARD; PRT; 2196 AA.  
AC Q9HDV6;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cell polarity protein mor2 (Morphological round protein 2).  
GN MOR2 OR SPBP19A11.04C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.  
RC STRAIN=972;  
RX MEDLINE=22220058; PubMed=12234926;  
RA Hirata D., Kishimoto N., Suda M., Sogabe Y., Nakagawa S., Yoshida Y.,  
RA Sakai K., Mizunuma M., Miyakawa T., Ishiguro J., Toda T.;  
RT "Fission yeast Mor2/Cps12, a protein similar to Drosophila Furry, is  
RT essential for cell morphogenesis and its mutation induces  
RT Wee1-dependent G2 delay."  
RL EMBO J. 21:4863-4874(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe."  
RL Nature 415:871-880(2002).  
CC -!- FUNCTION: Required for the maintenance of cell polarity. Has a  
CC role in localizing F-actin at the cell tips.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated; found

CC at the cell tips and at the site of septum formation.

CC -!- SIMILARITY: TO YEAST TAO3/PAG1.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AB093007; BAC20935.1; -.

DR EMBL; AL512495; CAC19754.1; -.

DR GeneDB SPombe; SPBP19A11.04c; -.

SQ SEQUENCE 2196 AA; 251195 MW; AAF2698DB00DE4CF CRC64;

Query Match 77.5%; Score 31; DB 1; Length 2196;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8

|||||

Db 2014 VFFAED 2019

#### RESULT 20

UT11\_ORYSA

ID UT11\_ORYSA STANDARD; PRT; 214 AA.

AC Q8S1Z1;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Probable U3 small nucleolar RNA-associated protein 11 (U3 snoRNA-  
DE associated protein 11).

GN P0468B07.17.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI\_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC

clone:P0468B07.";

RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: Involved in nucleolar processing of pre-18S ribosomal  
CC RNA (By similarity).

CC -!- SUBUNIT: Component of the ribosomal small subunit (SSU)  
CC processome (By similarity).

CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).

CC -!- SIMILARITY: Belongs to the UTP11 family.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AP003260; BAB89638.1; -.  
DR Gramene; Q8S1Z1; -.  
DR InterPro; IPR007144; Utp11.  
DR Pfam; PF03998; Utp11; 1.  
KW rRNA processing; Nuclear protein.  
SQ SEQUENCE 214 AA; 25321 MW; 4352134326922B43 CRC64;

Query Match 75.0%; Score 30; DB 1; Length 214;  
Best Local Similarity 75.0%; Pred. No. 20;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
| | : | | | |  
Db 130 KHVYFAED 137

#### RESULT 21

Y665\_METJA

ID Y665\_METJA STANDARD; PRT; 341 AA.

AC Q58079;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein MJ0665.

GN MJ0665.

OS Methanococcus jannaschii.

OC Archaea; Euryarchaeota; Methanococci; Methanococcales;

OC Methanocaldococcaceae; Methanocaldococcus.

OX NCBI\_TaxID=2190;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;

RX MEDLINE=96337999; PubMed=8688087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,

RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,

RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,

RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,

RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,

RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,

RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;

RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii.";

RL Science 273:1058-1073(1996).  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; U67513; AAB98656.1; -.  
DR PIR; A64383; A64383.  
DR TIGR; MJ0665; -.  
DR InterPro; IPR007152; DUF354.  
DR Pfam; PF04007; DUF354; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 341 AA; 38789 MW; DC678410BC847458 CRC64;

Query Match 75.0%; Score 30; DB 1; Length 341;  
Best Local Similarity 71.4%; Pred. No. 32;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
||:|:|  
Db 67 KLIFYAE 73

#### RESULT 22

##### GP85 HUMAN

ID GP85\_HUMAN STANDARD; PRT; 370 AA.  
AC Q9NPD1; Q9JHI6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Probable G protein-coupled receptor GPR85 (Super conserved receptor  
DE expressed in brain 2) (PKrCx1).  
GN GPR85 OR SREB2.  
OS Homo sapiens (Human),  
OS Mus musculus (Mouse), and  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606, 10090, 10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human, and Mouse; STRAIN=CD-1; TISSUE=Brain, and Fetal brain;  
RX MEDLINE=20435311; PubMed=10978537;  
RA Hellebrand S., Schaller H.C., Wittenberger T.;  
RT "The brain-specific G-protein coupled receptor GPR85 with identical  
RT protein sequence in man and mouse maps to human chromosome 7q31.";  
RL Biochim. Biophys. Acta 1493:269-272(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human, and Rat; TISSUE=Brain;  
RX MEDLINE=20294882; PubMed=10833454;  
RA Matsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,  
RA Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuichi K.;  
RT "An evolutionarily conserved G-protein coupled receptor family, SREB,  
RT expressed in the central nervous system.";  
RL Biochem. Biophys. Res. Commun. 272:576-582(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human;  
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,  
RA Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;  
RT "Genome-wide discovery and analysis of human seven transmembrane helix

RT receptor genes.";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Amygdala;  
 RA Poustka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Rat; STRAIN=Sprague-Dawley;  
 RA Kim H., Park S., Kang Y., Kim C., Jeon J.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Mouse; STRAIN=C57BL/6; TISSUE=Eye;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Orphan receptor.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- TISSUE SPECIFICITY: IN HUMAN HIGHLY EXPRESSED IN BRAIN AND TESTIS.  
 CC LOWER LEVELS IN SMALL INTESTINE, PLACENTA AND SPLEEN. IN BRAIN  
 CC REGIONS, DETECTED IN ALL REGIONS TESTED, BUT SOMEWHAT LOWER LEVELS  
 CC IN THE CORPUS CALLOSUM, MEDULLA AND SPINAL CORD. IN THE MOUSE  
 CC EXCLUSIVELY EXPRESSED IN BRAIN.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF250237; AAF79956.1; -.  
 DR EMBL; AB040800; BAA96646.1; -.  
 DR EMBL; AL161959; CAB82307.1; -.  
 DR EMBL; AF254416; AAF79959.1; -.

DR EMBL; AB065688; BAC05911.1; -.  
 DR EMBL; AB040803; BAA96649.1; -.  
 DR EMBL; AF203907; AAG42284.1; -.  
 DR EMBL; BC026975; AAH26975.1; -.  
 DR PIR; T47131; T47131.  
 DR Genew; HGNC:4536; GPR85.  
 DR MIM; 605188; -.  
 DR MGD; MGI:1927851; Gpr85.  
 DR GO; GO:0004930; F:G-protein coupled receptor activity; TAS.  
 DR GO; GO:0007165; P:signal transduction; TAS.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; FALSE\_NEG.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family.  
 FT DOMAIN 1 25 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 26 46 1 (POTENTIAL).  
 FT DOMAIN 47 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78 2 (POTENTIAL).  
 FT DOMAIN 79 96 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 97 117 3 (POTENTIAL).  
 FT DOMAIN 118 137 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 138 158 4 (POTENTIAL).  
 FT DOMAIN 159 188 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 189 209 5 (POTENTIAL).  
 FT DOMAIN 210 286 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 287 307 6 (POTENTIAL).  
 FT DOMAIN 308 313 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 314 334 7 (POTENTIAL).  
 FT DOMAIN 335 370 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 94 172 BY SIMILARITY.  
 FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 370 AA; 41994 MW; 7B67A39F6166AAEB CRC64;

Query Match 75.0%; Score 30; DB 1; Length 370;  
 Best Local Similarity 62.5%; Pred. No. 34;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||:|| |  
 Db 204 KLIFVVD 211

RESULT 23  
 GP85\_BRARE  
 ID GP85\_BRARE STANDARD; PRT; 371 AA.  
 AC Q9I919;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Super conserved receptor expressed in brain 2.  
 GN SREB2.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20294882; PubMed=10833454;  
 RA Matsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,  
 RA Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuichi K.;  
 RT "An evolutionarily conserved G-protein coupled receptor family, SREB,  
 RT expressed in the central nervous system.";  
 RL Biochem. Biophys. Res. Commun. 272:576-582(2000).  
 CC -!- FUNCTION: Orphan receptor.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB040805; BAA96651.1; -.  
 DR ZFIN; ZDB-GENE-000710-2; sreb2.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; FALSE\_NEG.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family.  
 FT DOMAIN 1 26 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 27 47 1 (POTENTIAL).  
 FT DOMAIN 48 58 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 59 79 2 (POTENTIAL).  
 FT DOMAIN 80 96 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 97 117 3 (POTENTIAL).  
 FT DOMAIN 118 138 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 139 159 4 (POTENTIAL).  
 FT DOMAIN 160 189 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 190 210 5 (POTENTIAL).  
 FT DOMAIN 211 287 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 288 308 6 (POTENTIAL).  
 FT DOMAIN 309 321 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 322 342 7 (POTENTIAL).  
 FT DOMAIN 343 371 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 95 173 BY SIMILARITY.  
 FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 371 AA; 41954 MW; F6F6175ED3A348C2 CRC64;

Query Match 75.0%; Score 30; DB 1; Length 371;  
 Best Local Similarity 62.5%; Pred. No. 34;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy            1 KLVFFAED 8  
             ||:|| |  
Db           205 KLIFVVD 212

RESULT 24

SYK METMP

ID SYK\_METMP            STANDARD;            PRT;    533 AA.  
AC O30522;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Lysyl-tRNA synthetase (EC 6.1.1.6) (Lysine--tRNA ligase) (LysRS).  
GN LYSS.  
OS Methanococcus maripaludis.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanococcaceae; Methanococcus.  
OX NCBI\_TaxID=39152;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-22.  
RC STRAIN=JJ;  
RX MEDLINE=98016282; PubMed=9353192;  
RA Ibba M., Morgan S., Curnow A.W., Pridmore D.R., Vothknecht U.C.,  
RA Gardner W., Lin W., Woese C.R., Soell D.;  
RT "A euryarchaeal lysyl-tRNA synthetase: resemblance to class I  
RT synthetases.";  
RL Science 278:1119-1122(1997).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=20570460; PubMed=11121028;  
RA Soell D., Becker H.D., Plateau P., Blanquet S., Ibba M.;  
RT "Context-dependent anticodon recognition by class I lysyl-tRNA  
RT synthetases.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:14224-14228(2000).  
CC -!- CATALYTIC ACTIVITY: ATP + L-lysine + tRNA(Lys) = AMP + diphosphate  
CC    + L-lysyl-tRNA(Lys).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- MISCELLANEOUS: Able to charge E.coli tRNA(Lys) in vitro.  
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF009824; AAB87410.1; -.  
DR PIR; T46975; T46975.  
DR HAMAP; MF\_00177; -; 1.  
DR InterPro; IPR002904; Lys\_tRNA-synt\_lc.  
DR InterPro; IPR008925; tRNA-synt\_bind.  
DR InterPro; IPR001412; tRNA-synt\_I.  
DR Pfam; PF01921; tRNA-synt\_1f; 1.  
DR TIGRFAMs; TIGR00467; lysS\_arch; 1.

DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 FT SITE 28 36 "HIGH" REGION.  
 FT SITE 278 282 "KMSKS" REGION.  
 FT CONFLICT 11 11 I -> L (IN REF. 1; AA SEQUENCE).  
 SQ SEQUENCE 533 AA; 61273 MW; 007FA868A57A0AC2 CRC64;

Query Match 75.0%; Score 30; DB 1; Length 533;  
 Best Local Similarity 62.5%; Pred. No. 49;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 ||:| |:|  
 Db 58 KLIFIADD 65

# RESULT 25

## CYA9\_HUMAN

ID CYA9\_HUMAN STANDARD; PRT; 1353 AA.  
 AC O60503; O60273; Q9BWT4; Q9UGP2;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Adenylate cyclase, type IX (EC 4.6.1.1) (ATP pyrophosphate-lyase)  
 DE (Adenylyl cyclase).  
 GN ADCY9 OR KIAA0520.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=98292499; PubMed=9628827;  
 RA Hacker B.M., Tomlinson J.E., Wayman G.A., Sultana R., Chan G.,  
 RA Villacres E., Distech C., Storm D.R.;  
 RT "Cloning, chromosomal mapping, and regulatory properties of the human  
 RT type 9 adenylyl cyclase (ADCY9).";  
 RL Genomics 50:97-104(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Paterson J.M., Smith S.M., Simpson J., Grace O.C., Bell J.E.,  
 RA Antoni F.A.;  
 RT "Cloning and characterisation of human adenylyl cyclase IX:  
 RT differential mRNA regulation and inhibition by Ca2+/calcineurin.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Toyota T., Yamada K., Meerabux J., Hattori E., Saito K.,  
 RA Yoshitsugu K., Shimizu H., Nankai M., Toru M., Detera-Wadleigh S.D.,  
 RA Yoshikawa T.;  
 RT "Mutation screening, case control study and transmission  
 RT disequilibrium analysis of adenylate cyclase type 9 (ADCY9) gene in  
 RT functional psychoses.";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE OF 141-1353 FROM N.A., AND REVISION TO 1154.

RC TISSUE=Brain;  
 RX MEDLINE=22158633; PubMed=12168954;  
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;  
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual  
 RT curation of 330 KIAA cDNA clones.";  
 RL DNA Res. 9:99-106(2002).  
 RN [5]  
 RP SEQUENCE OF 788-1353 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98290545; PubMed=9628581;  
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,  
 RA Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. IX.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:31-39(1998).  
 CC -!- FUNCTION: May play a fundamental role in situations where fine  
 CC interplay between intracellular calcium and cAMP determines the  
 CC cellular function. May be a physiologically relevant docking site  
 CC for calcineurin (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.  
 CC -!- COFACTOR: Binds 2 magnesium ions per subunit.  
 CC -!- ENZYME REGULATION: Insensitive to calcium/calmodulin, forskolin  
 CC and somatostatin. Stimulated by beta-adrenergic receptor  
 CC activation.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- DOMAIN: Composed of two homologous domains.  
 CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-4/guanylyl  
 CC cyclase family.  
 CC -!- SIMILARITY: Contains 1 FKBP-type PPIase domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF036927; AAC24201.1; -.  
 DR EMBL; AJ133123; CAB65084.1; -.  
 DR EMBL; AY028959; AAK29464.1; -.  
 DR EMBL; AY028949; AAK29464.1; JOINED.  
 DR EMBL; AY028950; AAK29464.1; JOINED.  
 DR EMBL; AY028951; AAK29464.1; JOINED.  
 DR EMBL; AY028952; AAK29464.1; JOINED.  
 DR EMBL; AY028953; AAK29464.1; JOINED.  
 DR EMBL; AY028954; AAK29464.1; JOINED.  
 DR EMBL; AY028955; AAK29464.1; JOINED.  
 DR EMBL; AY028956; AAK29464.1; JOINED.  
 DR EMBL; AY028957; AAK29464.1; JOINED.  
 DR EMBL; AB011092; BAA25446.2; -.  
 DR HSSP; P26769; 1AB8.  
 DR Genew; HGNC:240; ADCY9.  
 DR MIM; 603302; -.  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0004016; F:adenylate cyclase activity; TAS.

DR GO; GO:0007165; P:signal transduction; TAS.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR Pfam; PF00211; guanylate\_cyc; 2.  
 DR SMART; SM00044; CYCc; 2.  
 DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 2.  
 KW Lyase; cAMP biosynthesis; Transmembrane; Glycoprotein; Repeat;  
 KW Metal-binding; Magnesium; Isomerase; Rotamase.  
 FT DOMAIN 1 117 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 118 138 POTENTIAL.  
 FT DOMAIN 139 141 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 142 162 POTENTIAL.  
 FT DOMAIN 163 171 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 172 192 POTENTIAL.  
 FT DOMAIN 193 215 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 216 235 POTENTIAL.  
 FT DOMAIN 236 241 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 242 259 POTENTIAL.  
 FT DOMAIN 260 280 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 281 301 POTENTIAL.  
 FT DOMAIN 302 786 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 787 807 POTENTIAL.  
 FT DOMAIN 808 818 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 819 839 POTENTIAL.  
 FT DOMAIN 840 867 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 868 888 POTENTIAL.  
 FT DOMAIN 889 891 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 892 912 POTENTIAL.  
 FT DOMAIN 913 920 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 921 941 POTENTIAL.  
 FT DOMAIN 942 975 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 976 996 POTENTIAL.  
 FT DOMAIN 997 1353 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 504 611 PPIASE, FKBP-TYPE.  
 FT METAL 399 399 MAGNESIUM 1 AND 2 (BY SIMILARITY).  
 FT METAL 400 400 MAGNESIUM 2 (VIA CARBONYL OXYGEN) (BY  
 FT SIMILARITY).  
 FT METAL 443 443 MAGNESIUM 1 AND 2 (BY SIMILARITY).  
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 955 955 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 493 493 G -> R (IN REF. 2).  
 FT CONFLICT 884 884 V -> A (IN REF. 2).  
 FT CONFLICT 1154 1154 N -> S (IN REF. 2 AND 4).  
 FT CONFLICT 1253 1353 QHQLSISPDIRVQVDGSGIGRSPTDEIANLVPSVQYVDKTSL  
 FT GSDSSTQAKDAHLSPKRPWKEPVKAEERGRFGKAIEKDDCD  
 FT ETGIEEANELTKLNVSKSV -> APAVHLPRHPRGGWQHR  
 FT TVSHRRDCQPGAFPCVCGQDISGF (IN REF. 1).  
 FT CONFLICT 1308 1308 P -> R (IN REF. 2).  
 SQ SEQUENCE 1353 AA; 150699 MW; 4CBF051EA49B5B7B CRC64;

Query Match 75.0%; Score 30; DB 1; Length 1353;  
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ::||| ||

## RESULT 26

## CYA9\_MOUSE

ID CYA9\_MOUSE STANDARD; PRT; 1353 AA.  
AC P51830; Q61279;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Adenylate cyclase, type IX (EC 4.6.1.1) (ATP pyrophosphate-lyase)  
DE (Adenylyl cyclase) (Adenylyl cyclase type 10) (ACTP10).  
GN ADCY9.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c; TISSUE=Brain;  
RX MEDLINE=96278831; PubMed=8662814;  
RA Premont R.T., Matsuoka I., Mattei M.-G., Pouille Y., Defer N.,  
RA Hanoune J.;  
RT "Identification and characterization of a widely expressed form of  
RT adenylyl cyclase.";  
RL J. Biol. Chem. 271:13900-13907(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96024597; PubMed=7575502;  
RA Paterson J.M., Smith S.M., Harmar A.J., Antoni F.A.;  
RT "Control of a novel adenylyl cyclase by calcineurin.";  
RL Biochem. Biophys. Res. Commun. 214:1000-1008(1995).  
RN [3]  
RP SEQUENCE OF 1106-1193 FROM N.A.  
RX MEDLINE=95097788; PubMed=7528319;  
RA Premont R.T.;  
RT "Identification of adenylyl cyclases by amplification using  
RT degenerate primers.";  
RL Meth. Enzymol. 238:116-127(1994).  
CC -!- FUNCTION: May play a fundamental role in situations where fine  
CC interplay between intracellular calcium and cAMP determines the  
CC cellular function. May be a physiologically relevant docking site  
CC for calcineurin.  
CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.  
CC -!- COFACTOR: Binds 2 magnesium ions per subunit.  
CC -!- ENZYME REGULATION: Insensitive to calcium/calmodulin. Stimulated  
CC by magnesium, forskolin and mutationally activated G protein (GS)-  
CC alpha. Regulated by calcineurin.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: Found in decreasing order in skeletal muscle,  
CC heart, adrenal gland, ovary and brain; and to a lesser extent, in  
CC kidney, liver, testis, lung, thymus and spleen.  
CC -!- DOMAIN: Composed of two homologous domains.  
CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-4/guanylyl  
CC cyclase family.  
CC -!- SIMILARITY: Contains 1 FKBP-type PPIase domain.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U30602; AAC52603.1; -.  
 DR EMBL; Z50190; CAA90570.1; -.  
 DR PIR; JC4279; JC4279.  
 DR HSSP; P26769; 1AB8.  
 DR MGD; MGI:108450; Adcy9.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR Pfam; PF00211; guanylate\_cyc; 2.  
 DR SMART; SM00044; CYCc; 2.  
 DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 2.  
 KW Lyase; cAMP biosynthesis; Transmembrane; Glycoprotein; Repeat;  
 KW Metal-binding; Magnesium; Isomerase; Rotamase.  
 FT DOMAIN 1 117 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 118 138 POTENTIAL.  
 FT DOMAIN 139 141 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 142 162 POTENTIAL.  
 FT DOMAIN 163 171 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 172 192 POTENTIAL.  
 FT DOMAIN 193 215 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 216 235 POTENTIAL.  
 FT DOMAIN 236 241 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 242 259 POTENTIAL.  
 FT DOMAIN 260 280 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 281 301 POTENTIAL.  
 FT DOMAIN 302 786 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 787 807 POTENTIAL.  
 FT DOMAIN 808 818 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 819 839 POTENTIAL.  
 FT DOMAIN 840 867 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 868 888 POTENTIAL.  
 FT DOMAIN 889 891 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 892 912 POTENTIAL.  
 FT DOMAIN 913 920 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 921 941 POTENTIAL.  
 FT DOMAIN 942 975 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 976 996 POTENTIAL.  
 FT DOMAIN 997 1353 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 504 611 PPIASE, FKBP-TYPE.  
 FT METAL 399 399 MAGNESIUM 1 AND 2 (BY SIMILARITY).  
 FT METAL 400 400 MAGNESIUM 2 (VIA CARBONYL OXYGEN) (BY  
 FT SIMILARITY).  
 FT METAL 443 443 MAGNESIUM 1 AND 2 (BY SIMILARITY).  
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 955 955 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 1192 1192 N -> D (IN REF. 3).  
 FT CONFLICT 1305 1305 R -> H (IN REF. 2).  
 SQ SEQUENCE 1353 AA; 150953 MW; C65736A8304F689E CRC64;

Query Match 75.0%; Score 30; DB 1; Length 1353;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
::||| ||  
Db 840 RMVFFLED 847

RESULT 27

RL7A\_HALMA

ID RL7A\_HALMA STANDARD; PRT; 119 AA.  
AC P12743;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE 50S ribosomal protein L7Ae (HS6).  
GN RPL7AE.  
OS Haloarcula marismortui (Halobacterium marismortui).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Haloarcula.  
OX NCBI\_TaxID=2238;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=88055606; PubMed=3315748;  
RA Kimura J., Arndt E., Kimura M.;  
RT "Primary structures of three highly acidic ribosomal proteins S6, S12  
RT and S15 from the archaebacterium Halobacterium marismortui.";  
RL FEBS Lett. 224:65-70(1987).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS), AND REVISIONS TO 108-110.  
RC STRAIN=ATCC 43049;  
RX MEDLINE=20396344; PubMed=10937989;  
RA Ban N., Nissen P., Hansen J., Moore P.B., Steitz T.A.;  
RT "The complete atomic structure of the large ribosomal subunit at 2.4  
RT A resolution.";  
RL Science 289:905-920(2000).  
CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.  
CC -!- CAUTION: Was originally (Ref.1) thought to be a protein from the  
CC small subunit of the ribosome.  
DR PIR; S00182; R5HSS6.  
DR PDB; 1FFK; 26-SEP-01.  
DR PDB; 1K8A; 19-JUL-02.  
DR PDB; 1K9M; 19-JUL-02.  
DR PDB; 1KD1; 19-JUL-02.  
DR PDB; 1M1K; 23-AUG-02.  
DR PDB; 1M90; 06-SEP-02.  
DR HAMAP; MF\_00326; -; 1.  
DR InterPro; IPR004038; Ribosomal\_L7A.  
DR InterPro; IPR004037; Ribosomal\_L7Ae.  
DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
DR PRINTS; PR00881; L7ARS6FAMILY.  
DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.  
KW Ribosomal protein; 3D-structure.  
FT INIT\_MET 0 0  
FT CONFLICT 108 110 MISSING (IN REF. 1).  
SQ SEQUENCE 119 AA; 12591 MW; A132AA40519F45C5 CRC64;



Query Match 72.5%; Score 29; DB 1; Length 119;  
Best Local Similarity 75.0%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 :||| |||  
Db 46 ELVFVAED 53

RESULT 28

CME2\_BACSU

ID CME2\_BACSU STANDARD; PRT; 189 AA.  
AC P32393;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE ComE operon protein 2.  
GN COMEB OR COME2 OR BSU25580.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95058187; PubMed=7968523;  
RA Hahn J., Inamine G., Kozlov Y., Dubnau D.A.;  
RT "Characterization of comE, a late competence operon of Bacillus  
RT subtilis required for the binding and uptake of transforming DNA."  
RL Mol. Microbiol. 10:99-110(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / JH642;  
RX MEDLINE=97124195; PubMed=8969508;  
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,  
RA Kobayashi Y.;  
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of  
RT the Bacillus subtilis genome containing the skin element and many  
RT sporulation genes."  
RL Microbiology 142:3103-3111(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL Nature 390:249-256(1997).

CC -!- FUNCTION: DISPENSABLE FOR TRANSFORMABILITY.  
 CC -!- COFACTOR: Zinc (By similarity).  
 CC -!- SIMILARITY: Belongs to the cytidine and deoxycytidylate deaminases  
 CC family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L15202; AAC36906.1; -.  
 DR EMBL; D84432; BAA12453.1; -.  
 DR EMBL; Z99117; CAB14500.1; -.  
 DR PIR; S39864; S39864.  
 DR SubtiList; BG10481; comEB.  
 DR InterPro; IPR002125; dCMP/cyt\_deam.  
 DR Pfam; PF00383; dCMP\_cyt\_deam; 1.  
 DR PROSITE; PS00903; CYT\_DCMP\_DEAMINASES; 1.  
 KW Hydrolase; Zinc; Complete proteome.  
 FT METAL 70 70 ZINC (BY SIMILARITY).  
 FT METAL 98 98 ZINC (BY SIMILARITY).  
 FT METAL 101 101 ZINC (BY SIMILARITY).  
 SQ SEQUENCE 189 AA; 20969 MW; B6AC951966EE2D37 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 189;  
 Best Local Similarity 62.5%; Pred. No. 30;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 | |:|  
 Db 111 KTVVYAED 118

# RESULT 29

RSFA\_BACSU

ID RSFA\_BACSU STANDARD; PRT; 258 AA.  
 AC P39650;

DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Prespore specific transcriptional regulator rsfA.  
 GN RSFA OR IPA-92R OR BSU37620.  
 OS *Bacillus subtilis*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=95020537; PubMed=7934828;  
 RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,  
 RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,  
 RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,  
 RA Rapoport G., Danchin A.;  
 RT "Bacillus subtilis genome project: cloning and sequencing of the 97  
 RT kb region from 325 degrees to 333 degrees.";  
 RL Mol. Microbiol. 10:371-384(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL Nature 390:249-256(1997).  
 RN [3]  
 RP CHARACTERIZATION.  
 RC STRAIN=168;

RX MEDLINE=20096685; PubMed=10629188;  
 RA Wu L.J., Errington J.;  
 RT "Identification and characterization of a new prespore-specific  
 RT regulatory gene, rsfA, of *Bacillus subtilis*.";  
 RL J. Bacteriol. 182:418-424(2000).  
 CC -!- FUNCTION: Seems to improve the efficiency of sporulation by fine-  
 CC tuning the expression of genes in the sigma F regulon,  
 CC particularly the timing of their expression. Negatively regulates  
 CC spoIIR and its own synthesis.  
 CC -!- SUBCELLULAR LOCATION: Expressed in the forespore. Colocalizes with  
 CC DNA.  
 CC -!- DEVELOPMENTAL STAGE: Expressed during sporulation.  
 CC -!- INDUCTION: Expression is sigma F and sigma G-dependent.  
 CC -!- SIMILARITY: TO B.SUBTILIS YLBO.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X73124; CAA51648.1; -.  
 DR EMBL; Z99123; CAB15789.1; -.  
 DR PIR; S39747; S39747.  
 DR SubtiList; BG10638; rsfA.  
 KW Sporulation; Transcription regulation; Activator; Repressor;  
 KW DNA-binding; Complete proteome.  
 SQ SEQUENCE 258 AA; 29728 MW; 67C65A0483BF8B5C CRC64;

Query Match 72.5%; Score 29; DB 1; Length 258;  
 Best Local Similarity 75.0%; Pred. No. 41;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||| | ||  
 Db 231 KLVLFEE 238

#### RESULT 30

##### ERA\_BUCAP

ID ERA\_BUCAP STANDARD; PRT; 278 AA.  
 AC Q8K9R2;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE GTP-binding protein era homolog.  
 GN ERA OR BUSG248.  
 OS Buchnera aphidicola (subsp. Schizaphis graminum).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=98794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22084549; PubMed=12089438;  
 RA Tamas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,

RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;  
 RT "50 million years of genomic stasis in endosymbiotic bacteria.";  
 RL Science 296:2376-2379(2002).  
 CC -!- FUNCTION: Binds both GDP and GTP, has an intrinsic GTPase activity  
 CC and is essential for cell growth (By similarity).  
 CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding  
 CC proteins. Era subfamily.  
 CC -!- SIMILARITY: Contains 1 KH type-2 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE014101; AAM67807.1; -.  
 DR HAMAP; MF\_00367; -; 1.  
 DR InterPro; IPR005662; Era.  
 DR InterPro; IPR006073; GTP1\_OBG.  
 DR InterPro; IPR009019; KH\_prok.  
 DR InterPro; IPR004044; KH\_TYPE\_2.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR PRINTS; PR00326; GTP1OBG.  
 DR TIGRFAMs; TIGR00436; era; 1.  
 DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
 DR PROSITE; PS50823; KH\_TYPE\_2; 1.  
 KW GTP-binding; RNA-binding; Complete proteome.  
 FT NP\_BIND 15 22 GTP (POTENTIAL).  
 FT NP\_BIND 62 66 GTP (POTENTIAL).  
 FT NP\_BIND 117 120 GTP (POTENTIAL).  
 FT DOMAIN 199 276 KH TYPE-2.  
 SQ SEQUENCE 278 AA; 32487 MW; 05489832FAE85903 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 278;  
 Best Local Similarity 62.5%; Pred. No. 44;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 ||:|| |  
 Db 194 KLIFFLRD 201

RESULT 31  
 DMA\_HAEIN  
 ID DMA\_HAEIN STANDARD; PRT; 286 AA.  
 AC P44431;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE DNA adenine methylase (EC 2.1.1.72) (Deoxyadenosyl-methyltransferase)  
 DE (DNA adenine methyltransferase) (M.HindIV).  
 GN DAM OR HINDIVM OR HI0209.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.

OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT Rd.";  
 RL Science 269:496-512(1995).  
 CC -!- FUNCTION: THIS ENZYME METHYLATES DNA WITHIN THE SEQUENCE GATC.  
 CC DIRECTLY INVOLVED IN METHYL-DIRECTED DNA MISMATCH REPAIR (BY  
 CC SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-  
 CC adenosyl-L-homocysteine + DNA 6-methylaminopurine.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U32705; AAC21877.1; -.  
 DR PIR; H64054; H64054.  
 DR HSSP; P04043; 2DPM.  
 DR REBASE; 1152; M.HindIV.  
 DR TIGR; HI0209; -.  
 DR InterPro; IPR002294; D12N6\_mtfrase.  
 DR InterPro; IPR002052; N6\_Mtase.  
 DR Pfam; PF02086; MethyltransfD12; 1.  
 DR PRINTS; PR00505; D12N6MTFRASE.  
 DR TIGRFAMS; TIGR00571; dam; 1.  
 DR PROSITE; PS00092; N6\_MTASE; 1.  
 KW Transferase; Methyltransferase; DNA replication; Complete proteome.  
 SQ SEQUENCE 286 AA; 33218 MW; BE39C5725AE35DEE CRC64;

Query Match 72.5%; Score 29; DB 1; Length 286;  
 Best Local Similarity 62.5%; Pred. No. 45;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 | :|||:|  
 Db 90 KPIFFADD 97

RESULT 32  
 HST2\_YEAST  
 ID HST2\_YEAST STANDARD; PRT; 357 AA.

AC P53686;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE NAD-dependent deacetylase HST2 (EC 3.5.1.-) (Homologous to SIR2  
DE protein 2).  
GN HST2 OR YPL015C OR LPA2C.  
OS *Saccharomyces cerevisiae* (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / YPH1;  
RX MEDLINE=96101589; PubMed=7498786;  
RA Brachmann C.B., Sherman J.M., Devine S.E., Cameron E.E., Pillus L.,  
RA Boeke J.D.;  
RT "The SIR2 gene family, conserved from bacteria to humans, functions  
RT in silencing, cell cycle progression, and chromosome stability.";  
RL Genes Dev. 9:2888-2902(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX MEDLINE=97313271; PubMed=9169875;  
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,  
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,  
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,  
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,  
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,  
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,  
RA Hunicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,  
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,  
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,  
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,  
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,  
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,  
RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,  
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;  
RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XVI.";  
RL Nature 387:103-105(1997).  
CC -!- FUNCTION: NAD-dependent histone deacetylase, which could function  
CC in telomeric silencing, cell cycle progression and chromosome  
CC stability.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + an acetylprotein = nicotinamide + O-  
CC acetyl-ADP-ribose + a protein.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -!- SIMILARITY: Belongs to the sirtuin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; U39063; AAA81035.1; -.  
 DR EMBL; U33335; AAB68090.1; -.  
 DR PIR; S59678; S59678.  
 DR GermOnline; 143997; -.  
 DR SGD; S0005936; HST2.  
 DR GO; GO:0005737; C:cytoplasm; IDA.  
 DR GO; GO:0017136; F:NAD-dependent histone deacetylase activity; IDA.  
 DR GO; GO:0006348; P:chromatin silencing at telomere; IGI.  
 DR InterPro; IPR003000; SIR2.  
 DR Pfam; PF02146; SIR2; 1.  
 DR PROSITE; PS50305; SIRTUIN; 1.  
 KW Hydrolase; NAD; Transcription regulation; Repressor; Nuclear protein;  
 KW Metal-binding; Zinc.  
 FT DOMAIN 13 286 DEACETYLASE SIRTUIN-TYPE.  
 FT ACT\_SITE 135 135 BY SIMILARITY.  
 FT METAL 143 143 ZINC (BY SIMILARITY).  
 FT METAL 146 146 ZINC (BY SIMILARITY).  
 FT METAL 170 170 ZINC (BY SIMILARITY).  
 FT METAL 173 173 ZINC (BY SIMILARITY).  
 SQ SEQUENCE 357 AA; 39979 MW; ED281E5B8241A4D0 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 357;  
 Best Local Similarity 71.4%; Pred. No. 56;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
 :||| ||  
 Db 181 IVFFGED 187

# RESULT 33

## RECJ\_METJA

ID RECJ\_METJA STANDARD; PRT; 469 AA.  
 AC Q58387;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Single-stranded-DNA-specific exonuclease recJ (EC 3.1.-.-).  
 GN RECJ OR MJ0977.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus



RT jannaschii.";  
 RL Science 273:1058-1073(1996).  
 RN [2]  
 RP ACTIVITY.  
 RX MEDLINE=20100748; PubMed=10633092;  
 RA Rajman L.A., Lovett S.T.;  
 RT "A thermostable single-strand DNase from Methanococcus jannaschii  
 RT related to the RecJ recombination and repair exonuclease from  
 RT Escherichia coli.";  
 RL J. Bacteriol. 182:607-612(2000).  
 CC -!- FUNCTION: SINGLE-STRANDED-DNA-SPECIFIC EXONUCLEASE.  
 CC -!- SIMILARITY: TO M.JANNASCHII MJ0831 AND MJ1071.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U67540; AAB98980.1; -.  
 DR PIR; A64422; A64422.  
 DR TIGR; MJ0977; -.  
 DR InterPro; IPR003156; DHHA1.  
 DR InterPro; IPR001667; Ppesterase.  
 DR Pfam; PF01368; DHH; 1.  
 DR Pfam; PF02272; DHHA1; 1.  
 KW Hydrolase; Nuclease; Exonuclease; Complete proteome.  
 SQ SEQUENCE 469 AA; 53847 MW; 037110C6D53B7554 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 469;  
 Best Local Similarity 75.0%; Pred. No. 73;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||: ||||  
 Db 417 KLLCFAED 424

#### RESULT 34

##### RNF9\_PANTR

ID RNF9\_PANTR STANDARD; PRT; 481 AA.  
 AC Q7YR32;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE RING finger protein 9 (B30-RING finger protein) (Tripartite motif  
 DE protein 10).  
 GN TRIM10 OR RNF9 OR RFB30.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22709134; PubMed=12799463;

RA Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenari A.,  
 RA Yamagata T., Kulski J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,  
 RA Yamazaki M., Tashiro H., Iwamoto C., Umehara Y., Imanishi T.,  
 RA Meyer A., Ikeo K., Gojobori T., Bahram S., Inoko H.;  
 RT "Comparative sequencing of human and chimpanzee MHC class I regions  
 RT unveils insertions/deletions as the major path to genomic  
 RT divergence.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713(2003).  
 CC -!- FUNCTION: Seems to play an important role in erythropoiesis (By  
 CC similarity).  
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -!- SIMILARITY: Contains 1 B box-type zinc finger.  
 CC -!- SIMILARITY: Contains 1 SPRY domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB100086; BAC78185.1; -.  
 DR InterPro; IPR001870; B302.  
 DR InterPro; IPR006574; PRY.  
 DR InterPro; IPR003877; SPRY\_receptor.  
 DR InterPro; IPR000315; Znf\_Bbox.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00622; SPRY; 1.  
 DR Pfam; PF00643; zf-B\_box; 1.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00336; BBOX; 1.  
 DR SMART; SM00589; PRY; 1.  
 DR SMART; SM00184; RING; 1.  
 DR SMART; SM00449; SPRY; 1.  
 DR PROSITE; PS50119; ZF\_BBOX; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Zinc-finger; Coiled coil.  
 FT ZN\_FING 16 61 RING-TYPE.  
 FT ZN\_FING 94 135 B BOX-TYPE.  
 FT DOMAIN 142 177 COILED COIL (POTENTIAL).  
 FT DOMAIN 362 481 SPRY.  
 SQ SEQUENCE 481 AA; 55112 MW; 67FACDDD78DEDE52 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 481;  
 Best Local Similarity 62.5%; Pred. No. 75;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |: || ||  
 Db 105 KIYFFCED 112

RESULT 35  
 RNF9\_HUMAN  
 ID RNF9\_HUMAN STANDARD; PRT; 482 AA.

AC Q9UDY6; Q96QB6; Q9C023; Q9C024;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE RING finger protein 9 (B30-RING finger protein) (Tripartite motif  
 DE protein 10).  
 GN TRIM10 OR RNF9 OR RFB30.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
 RC TISSUE=Testis;  
 RX MEDLINE=97419168; PubMed=9271628;  
 RA Henry J., Ribouchon M.-T., Depetris D., Mattei M.-G., Offer C.,  
 RA Tazi-Ahnini R., Pontarotti P.;  
 RT "Cloning, structural analysis, and mapping of B30 and B7 multigenic  
 RT families to the major histocompatibility complex (MHC) and other  
 RT chromosomal regions.";  
 RL Immunogenetics 46:383-395(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).  
 RX MEDLINE=21231161; PubMed=11331580;  
 RA Reymond A., Meroni G., Fantozzi A., Merla G., Cairo S., Luzi L.,  
 RA Riganelli D., Zanaria E., Messali S., Cainarca S., Guffanti A.,  
 RA Minucci S., Pelicci P.G., Ballabio A.;  
 RT "The tripartite motif family identifies cell compartments.";  
 RL EMBO J. 20:2140-2151(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
 RA Shiina S., Tamiya G., Oka A., Inoko H.;  
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Seems to play an important role in erythropoiesis (By  
 CC similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Alpha;  
 CC IsoId=Q9UDY6-1; Sequence=Displayed;  
 CC Name=Beta;  
 CC IsoId=Q9UDY6-2; Sequence=VSP\_005748;  
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -!- SIMILARITY: Contains 1 B box-type zinc finger.  
 CC -!- SIMILARITY: Contains 1 SPRY domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y07829; CAB52384.1; -.  
 DR EMBL; AF220122; AAG53495.1; -.  
 DR EMBL; AF220123; AAG53496.1; -.

DR EMBL; AP000517; BAB63332.1; -.  
 DR Genew; HGNC:10072; TRIM10.  
 DR MIM; 605701; -.  
 DR GO; GO:0005622; C:intracellular; NAS.  
 DR GO; GO:0008270; F:zinc ion binding; NAS.  
 DR GO; GO:0030097; P:hemopoiesis; NAS.  
 DR InterPro; IPR001870; B302.  
 DR InterPro; IPR006574; PRY.  
 DR InterPro; IPR003877; SPRY\_receptor.  
 DR InterPro; IPR000315; Znf\_Bbox.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00622; SPRY; 1.  
 DR Pfam; PF00643; zf-B\_box; 1.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00336; BBOX; 1.  
 DR SMART; SM00589; PRY; 1.  
 DR SMART; SM00184; RING; 1.  
 DR SMART; SM00449; SPRY; 1.  
 DR PROSITE; PS50119; ZF\_BBOX; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Zinc-finger; Coiled coil; Alternative splicing.  
 FT ZN\_FING 16 61 RING-TYPE.  
 FT ZN\_FING 94 136 B\_BOX-TYPE.  
 FT DOMAIN 143 178 COILED COIL (POTENTIAL).  
 FT DOMAIN 363 482 SPRY.  
 FT VARSPLIC 370 482 VSIDLAHGASCTVGVVSEDVQRKGELRLRPEEGVWAVRLAW  
 FT GFVSALGSFPTRLTLKEQPRQVRVSLDYEVGWVTFITNAVTR  
 FT EPIYTFTASFTRKVIPIFFGLWGRGSSFSLS -> WMARVP  
 FT GDSSGCGFCSPPSVLGTEVAA (in isoform Beta).  
 FT /FTId=VSP\_005748.  
 FT CONFLICT 104 105 SR -> G (IN REF. 1).  
 FT CONFLICT 104 105 SR -> E (IN REF. 3).  
 FT CONFLICT 168 168 K -> E (IN REF. 1 AND 3).  
 FT CONFLICT 337 337 S -> Y (IN REF. 2; AAG53496 AND 3).  
 FT CONFLICT 378 378 A -> G (IN REF. 3).  
 SQ SEQUENCE 482 AA; 55087 MW; 19354750602CD42C CRC64;

Query Match 72.5%; Score 29; DB 1; Length 482;  
 Best Local Similarity 62.5%; Pred. No. 75;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |: || ||  
 Db 106 KIYFFCED 113

# RESULT 36

## ELM1 ASPFU

ID ELM1 ASPFU STANDARD; PRT; 634 AA.  
 AC P46074;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Extracellular elastinolytic metalloproteinase precursor (EC 3.4.24.-).  
 GN MEP.  
 OS Aspergillus fumigatus (Sartorya fumigata).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 OX NCBI\_TaxID=5085;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 530-542.  
 RC STRAIN=Isolate 13;  
 RX MEDLINE=95012603; PubMed=7927676;  
 RA Sirakova T.D., Markaryan A., Kolattukudy P.E.;  
 RT "Molecular cloning and sequencing of the cDNA and gene for a novel  
 RT elastinolytic metalloproteinase from Aspergillus fumigatus and its  
 RT expression in Escherichia coli.";  
 RL Infect. Immun. 62:4208-4218(1994).

RN [2]  
 RP SEQUENCE OF 246-258.  
 RX MEDLINE=94245315; PubMed=8188335;  
 RA Markaryan A., Morozova I., Yu H., Kolattukudy P.E.;  
 RT "Purification and characterization of an elastinolytic  
 RT metalloprotease from Aspergillus fumigatus and immunoelectron  
 RT microscopic evidence of secretion of this enzyme by the fungus  
 RT invading the murine lung.";  
 RL Infect. Immun. 62:2149-2157(1994).  
 CC -!- FUNCTION: Catalyzes the hydrolysis of elastin.  
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -!- SIMILARITY: Belongs to peptidase family M36.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L29566; AAB07708.1; -.  
 DR MEROPS; M36.001; -.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001842; Peptidase\_M36.  
 DR Pfam; PF02128; Peptidase\_M36; 1.  
 DR PRINTS; PR00999; FUNGALYSIN.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Hydrolase; Metalloprotease; Zinc; Signal; Zymogen.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT PROPEP 19 245  
 FT CHAIN 246 634 EXTRACELLULAR ELASTINOLYTIC  
 FT METALLOPROTEINASE.  
 FT METAL 429 429 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 430 430 BY SIMILARITY.  
 FT METAL 433 433 ZINC (CATALYTIC) (BY SIMILARITY).  
 SQ SEQUENCE 634 AA; 68726 MW; B94E274BF767F911 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 634;  
 Best Local Similarity 62.5%; Pred. No. 98;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||:| :|  
 Db 198 KLVYFVKD 205

RESULT 37

ELM2\_ASPFU

ID ELM2\_ASPFU STANDARD; PRT; 634 AA.  
AC P46075;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Extracellular elastinolytic metalloproteinase precursor (EC 3.4.24.-).  
GN MEP.  
OS *Aspergillus fumigatus* (Sartorya fumigata).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.  
OX NCBI\_TaxID=5085;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=DELTA18;  
RX MEDLINE=95231298; PubMed=7715453;  
RA Jatton-Ogay K., Paris S., Huerre M., Quadroni M., Falchetto R.,  
RA Togni G., Latge J.-P., Monod M.;  
RT "Cloning and disruption of the gene encoding an extracellular  
RT metalloprotease of *Aspergillus fumigatus*."  
RL Mol. Microbiol. 14:917-928(1994).  
RN [2]  
RP REVISIONS.  
RC STRAIN=DELTA18;  
RA Sanglard D.;  
RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.  
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -!- SIMILARITY: Belongs to peptidase family M36.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z30424; CAA83015.1; -.  
DR MEROPS; M36.001; -.  
DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
DR InterPro; IPR001842; Peptidase\_M36.  
DR Pfam; PF02128; Peptidase\_M36; 1.  
DR PRINTS; PR00999; FUNGALYSIN.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
KW Hydrolase; Metalloprotease; Zinc; Signal; Zymogen.  
FT SIGNAL 1 18 POTENTIAL.  
FT PROPEP 19 245  
FT CHAIN 246 634 EXTRACELLULAR ELASTINOLYTIC  
FT METALLOPROTEINASE.  
FT METAL 429 429 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 430 430 BY SIMILARITY.  
FT METAL 433 433 ZINC (CATALYTIC) (BY SIMILARITY).  
SQ SEQUENCE 634 AA; 68718 MW; 7BA5A9B70DE184BE CRC64;

Query Match 72.5%; Score 29; DB 1; Length 634;  
Best Local Similarity 62.5%; Pred. No. 98;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||:| :|  
Db 198 KLVYFVKD 205

RESULT 38

KCB2\_CANFA

ID KCB2\_CANFA STANDARD; PRT; 809 AA.

AC Q95167;

DT 15-JUL-1999 (Rel. 38, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Potassium voltage-gated channel subfamily B member 2 (Potassium  
channel Kv2.2) (Fragment).

GN KCNB2.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI\_TaxID=9615;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98275219; PubMed=9612272;

RA Schmalz F., Kinsella J., Koh S.D., Vogalis F., Schneider A.,

RA Flynn E.R., Kenyon J.L., Horowitz B.;

RT "Molecular identification of a component of delayed rectifier current  
in gastrointestinal smooth muscles.";

RL Am. J. Physiol. 274:G901-G911(1998).

RN [2]

RP CONCEPTUAL TRANSLATION.

RA Sundaram S.;

RL Unpublished observations (OCT-2002).

CC -!- FUNCTION: Mediates the voltage-dependent potassium ion  
CC permeability of excitable membranes. Channels open or close in  
CC response to the voltage difference across the membrane, letting K+  
CC ions pass in accordance with their electrochemical gradient.

CC -!- SUBUNIT: Heteromultimer with KCNS1, KCNS2 and KCNS3 (By  
CC similarity).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is  
CC characterized by a series of positively charged amino acids at  
CC every third position.

CC -!- DOMAIN: The tail may be important in modulation of channel  
CC activity and/or targeting of the channel to specific subcellular  
CC compartments.

CC -!- PTM: Phosphorylated (Potential).

CC -!- SIMILARITY: Belongs to the potassium channel family. B (Shab)  
CC subfamily.

CC -!- CAUTION: This is a conceptual translation, a frameshift was  
CC introduced in position 768 to restore the similarity with the  
CC other orthologs.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U69963; AAB08432.1; ALT\_FRAME.  
DR HSSP; Q54397; 1BL8.  
DR InterPro; IPR000210; BTB\_POZ.  
DR InterPro; IPR005821; Ion\_trans.  
DR InterPro; IPR001622; K+channel\_pore.  
DR InterPro; IPR003091; K\_channel.  
DR InterPro; IPR003131; K\_tetra.  
DR InterPro; IPR005826; Kv22channel.  
DR InterPro; IPR003968; Kv\_channel.  
DR InterPro; IPR005820; M+channel\_nlg.  
DR InterPro; IPR003973; Shab\_channel.  
DR Pfam; PF00520; ion\_trans; 1.  
DR Pfam; PF02214; K\_tetra; 1.  
DR Pfam; PF03521; Kv2channel; 1.  
DR PRINTS; PR00169; KCHANNEL.  
DR PRINTS; PR01515; KV22CHANNEL.  
DR PRINTS; PR01491; KVCHANNEL.  
DR PRINTS; PR01495; SHABCHANNEL.  
DR SMART; SM00225; BTB; 1.  
KW Transport; Ion transport; Ionic channel; Voltage-gated channel;  
KW Potassium channel; Potassium; Potassium transport; Transmembrane;  
KW Phosphorylation; Glycoprotein; Multigene family.  
FT DOMAIN 1 190 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 191 212 SEGMENT S1 (POTENTIAL).  
FT TRANSMEM 233 254 SEGMENT S2 (POTENTIAL).  
FT DOMAIN 255 264 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 265 286 SEGMENT S3 (POTENTIAL).  
FT TRANSMEM 299 320 SEGMENT S4 (POTENTIAL).  
FT DOMAIN 321 334 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 335 356 SEGMENT S5 (POTENTIAL).  
FT DOMAIN 369 389 SEGMENT H5 (PORE-FORMING) (POTENTIAL).  
FT TRANSMEM 397 418 SEGMENT S6 (POTENTIAL).  
FT DOMAIN 419 >809 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT NON\_TER 809 809  
SQ SEQUENCE 809 AA; 90602 MW; E1FF56334F3DA9C8 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 809;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAE 7  
|||||  
Db 352 LVFFAE 357

#### RESULT 39

KCB1\_MOUSE

ID KCB1\_MOUSE STANDARD; PRT; 857 AA.

AC Q03717;

DT 28-FEB-2003 (Rel. 41, Created)



DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Potassium voltage-gated channel subfamily B member 1 (Potassium  
 DE channel Kv2.1) (mShab).  
 GN KCNB1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=91162315; PubMed=2002364;  
 RA Pak M.D., Covarrubias M., Ratcliffe A., Salkoff L.;  
 RT "A mouse brain homolog of the Drosophila Shab K<sup>+</sup> channel with  
 RT conserved delayed-rectifier properties.";  
 RL J. Neurosci. 11:869-880(1991).  
 CC -!- FUNCTION: Mediates the voltage-dependent potassium ion  
 CC permeability of excitable membranes. Channels open or close in  
 CC response to the voltage difference across the membrane, letting K<sup>+</sup>  
 CC ions pass in accordance with their electrochemical gradient.  
 CC -!- SUBUNIT: Heteromultimer with KCNG2, KCNG3, KCNG4, KCNS1, KCNS2,  
 CC KCNS3 and KCNV2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is  
 CC characterized by a series of positively charged amino acids at  
 CC every third position.  
 CC -!- DOMAIN: The tail may be important in modulation of channel  
 CC activity and/or targeting of the channel to specific subcellular  
 CC compartments.  
 CC -!- PTM: Phosphorylated on serine residues in the C-terminal  
 CC cytoplasmic tail (By similarity).  
 CC -!- MISCELLANEOUS: Inhibited by tetraethylammonium chloride (TEA).  
 CC -!- SIMILARITY: Belongs to the potassium channel family. B (Shab)  
 CC subfamily.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M64228; AAA40112.1; -.  
 DR PIR; I56529; I56529.  
 DR HSSP; Q54397; 1BL8.  
 DR MGD; MGI:96666; Kcnb1.  
 DR InterPro; IPR000210; BTB\_POZ.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR001622; K+channel\_pore.  
 DR InterPro; IPR003091; K\_channel.  
 DR InterPro; IPR003131; K\_tetra.  
 DR InterPro; IPR004350; Kv21channel.  
 DR InterPro; IPR003968; Kv\_channel.  
 DR InterPro; IPR005820; M+channel\_nlg.  
 DR InterPro; IPR003973; Shab\_channel.

DR Pfam; PF00520; ion\_trans; 1.  
 DR Pfam; PF02214; K\_tetra; 1.  
 DR Pfam; PF03521; Kv2channel; 1.  
 DR PRINTS; PR00169; KCHANNEL.  
 DR PRINTS; PR01514; KV21CHANNEL.  
 DR PRINTS; PR01491; KVCHANNEL.  
 DR PRINTS; PR01495; SHABCHANNEL.  
 DR SMART; SM00225; BTB; 1.  
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;  
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;  
 KW Phosphorylation; Multigene family.  
 FT DOMAIN 1 185 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 186 206 SEGMENT S1 (POTENTIAL).  
 FT TRANSMEM 229 249 SEGMENT S2 (POTENTIAL).  
 FT DOMAIN 250 259 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 260 280 SEGMENT S3 (POTENTIAL).  
 FT TRANSMEM 295 315 SEGMENT S4 (POTENTIAL).  
 FT DOMAIN 316 330 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 331 351 SEGMENT S5 (POTENTIAL).  
 FT DOMAIN 365 385 SEGMENT H5 (PORE-FORMING) (POTENTIAL).  
 FT TRANSMEM 393 413 SEGMENT S6 (POTENTIAL).  
 FT DOMAIN 414 857 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 517 520 POLY-SER.  
 SQ SEQUENCE 857 AA; 95692 MW; 772F1C42C25EF3A5 CRC64;  
  
 Query Match 72.5%; Score 29; DB 1; Length 857;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 2 LVFFAE 7  
 |||||  
 Db 348 LVFFAE 353

RESULT 40  
 KCB1\_RAT  
 ID KCB1\_RAT STANDARD; PRT; 857 AA.  
 AC P15387;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Potassium voltage-gated channel subfamily B member 1 (Potassium  
 DE channel Kv2.1) (DRK1).  
 GN KCNB1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE OF 4-857 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=89365157; PubMed=2770868;  
 RA Frech G.C., Vandongen A.M.J., Schuster G., Brown A.M., Joho R.H.;  
 RT "A novel potassium channel with delayed rectifier properties isolated  
 RT from rat brain by expression cloning."  
 RL Nature 340:642-645(1989).  
 RN [2]

RP REVISIONS.  
 RA Frech G.C.;  
 RL Submitted (FEB-1990) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-575 FROM N.A.  
 RX MEDLINE=92156897; PubMed=1740690;  
 RA Drewe J.A., Verma S., Frech G.C., Joho R.H.;  
 RT "Distinct spatial and temporal expression patterns of K<sup>+</sup> channel mRNAs  
 RT from different subfamilies.";  
 RL J. Neurosci. 12:538-548(1992).  
 RN [4]  
 RP PHOSPHORYLATION, AND MUTAGENESIS OF SER-444 AND SER-496.  
 RX MEDLINE=98016256; PubMed=9351973;  
 RA Murakoshi H., Shi G., Scannevin R.H., Trimmer J.S.;  
 RT "Phosphorylation of the Kv2.1 K<sup>+</sup> channel alters voltage-dependent  
 RT activation.";  
 RL Mol. Pharmacol. 52:821-828(1997).  
 CC -!- FUNCTION: Mediates the voltage-dependent potassium ion  
 CC permeability of excitable membranes. Channels open or close in  
 CC response to the voltage difference across the membrane, letting K<sup>+</sup>  
 CC ions pass in accordance with their electrochemical gradient.  
 CC -!- SUBUNIT: Heteromultimer with KCNG2, KCNG3, KCNG4, KCNS1, KCNS2,  
 CC KCNS3 and KCNV2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: In the brain, the greatest density occurs in  
 CC the cerebral cortex, followed by the hippocampus, cerebellum, and  
 CC olfactory bulb. In peripheral tissues it is most prominent in  
 CC retina and kidney. Also present in cardiac muscle tissue of the  
 CC atrium and ventricle and in skeletal muscle.  
 CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is  
 CC characterized by a series of positively charged amino acids at  
 CC every third position.  
 CC -!- DOMAIN: The tail may be important in modulation of channel  
 CC activity and/or targeting of the channel to specific subcellular  
 CC compartments.  
 CC -!- PTM: Phosphorylated on serine residues in the C-terminal  
 CC cytoplasmic tail, essentially between residues 667 and 853.  
 CC -!- SIMILARITY: Belongs to the potassium channel family. B (Shab)  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X16476; CAA34497.1; ALT\_INIT.  
 DR PIR; S05448; CHRTD1.  
 DR HSSP; Q54397; 1BL8.  
 DR InterPro; IPR000210; BTB\_POZ.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR001622; K\_channel\_pore.  
 DR InterPro; IPR003091; K\_channel.  
 DR InterPro; IPR003131; K\_tetra.  
 DR InterPro; IPR004350; Kv21channel.

DR InterPro; IPR003968; Kv\_channel.  
 DR InterPro; IPR005820; M+channel\_nlg.  
 DR InterPro; IPR003973; Shab\_channel.  
 DR Pfam; PF00520; ion\_trans; 1.  
 DR Pfam; PF02214; K\_tetra; 1.  
 DR Pfam; PF03521; Kv2channel; 1.  
 DR PRINTS; PR00169; KCHANNEL.  
 DR PRINTS; PR01514; KV21CHANNEL.  
 DR PRINTS; PR01491; KVCHANNEL.  
 DR PRINTS; PR01495; SHABCHANNEL.  
 DR SMART; SM00225; BTB; 1.  
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;  
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;  
 KW Phosphorylation; Multigene family.  
 FT DOMAIN 1 186 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 187 208 SEGMENT S1 (POTENTIAL).  
 FT TRANSMEM 229 250 SEGMENT S2 (POTENTIAL).  
 FT DOMAIN 251 260 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 261 282 SEGMENT S3 (POTENTIAL).  
 FT TRANSMEM 295 316 SEGMENT S4 (POTENTIAL).  
 FT DOMAIN 317 330 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 331 352 SEGMENT S5 (POTENTIAL).  
 FT DOMAIN 365 385 SEGMENT H5 (PORE-FORMING) (POTENTIAL).  
 FT TRANSMEM 393 414 SEGMENT S6 (POTENTIAL).  
 FT DOMAIN 415 857 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 517 520 POLY-SER.  
 FT DOMAIN 700 705 POLY-ALA.  
 FT MUTAGEN 444 444 S->A: NO EFFECT ON PHOSPHORYLATION.  
 FT MUTAGEN 496 496 S->A: NO EFFECT ON PHOSPHORYLATION.  
 SQ SEQUENCE 857 AA; 95637 MW; B3C5B0839AB15FD0 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 857;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAE 7  
 |||||  
 Db 348 LVFFAE 353

# RESULT 41

## KCB1\_HUMAN

ID KCB1\_HUMAN STANDARD; PRT; 858 AA.  
 AC Q14721; Q14193;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Potassium voltage-gated channel subfamily B member 1 (Potassium  
 channel Kv2.1) (h-DRK1).  
 GN KCNB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94363205; PubMed=8081723;

RA Albrecht B., Lorra C., Stocker K., Pongs O.;  
 RT "Cloning and characterization of a human delayed rectifier potassium  
 RT channel gene.";  
 RL Recept. Channels 1:99-110(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ikeda S.R., Soler F., Zuhlke R.D., Joho R.H., Lewis D.L.;  
 RL Submitted (JAN-1993) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lens epithelium;  
 RA Rae J.L., Shepard A.R.;  
 RT "Identification of potassium channels in human lens epithelium.";  
 RL (In) Civan M.M. (eds.);  
 RL The eye's aqueous humor-from secretion to glaucoma, pp.69-104,  
 RL Academic Press, San Diego (1998).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Lehvaslaiho M.H., Liversha M., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,  
 RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 CC -!- FUNCTION: Mediates the voltage-dependent potassium ion  
 CC permeability of excitable membranes. Channels open or close in  
 CC response to the voltage difference across the membrane, letting K+  
 CC ions pass in accordance with their electrochemical gradient.  
 CC -!- SUBUNIT: Heteromultimer with KCNG2, KCNG3, KCNG4, KCNS1, KCNS2,  
 CC KCNS3 and KCNV2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is  
 CC characterized by a series of positively charged amino acids at  
 CC every third position.  
 CC -!- DOMAIN: The tail may be important in modulation of channel

CC activity and/or targeting of the channel to specific subcellular  
 CC compartments.  
 CC -!- PTM: Phosphorylated on serine residues in the C-terminal  
 CC cytoplasmic tail (By similarity).  
 CC -!- SIMILARITY: Belongs to the potassium channel family. B (Shab)  
 CC subfamily.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X68302; CAA48374.1; -.  
 DR EMBL; I02840; AAA36156.1; ALT\_INIT.  
 DR EMBL; AF026005; AAB88808.1; -.  
 DR EMBL; AL035685; CAB89417.1; -.  
 DR PIR; S31761; S31761.  
 DR HSSP; Q54397; 1BL8.  
 DR Genew; HGNC:6231; KCNB1.  
 DR MIM; 600397; -.  
 DR InterPro; IPR000210; BTB\_POZ.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR001622; K+channel\_pore.  
 DR InterPro; IPR003091; K\_channel.  
 DR InterPro; IPR003131; K\_tetra.  
 DR InterPro; IPR004350; Kv21channel.  
 DR InterPro; IPR003968; Kv\_channel.  
 DR InterPro; IPR005820; M+channel\_nlg.  
 DR InterPro; IPR003973; Shab\_channel.  
 DR Pfam; PF00520; ion\_trans; 1.  
 DR Pfam; PF02214; K\_tetra; 1.  
 DR Pfam; PF03521; Kv2channel; 1.  
 DR PRINTS; PR00169; KCHANNEL.  
 DR PRINTS; PR01514; KV21CHANNEL.  
 DR PRINTS; PR01491; KVCHANNEL.  
 DR PRINTS; PR01495; SHABCHANNEL.  
 DR SMART; SM00225; BTB; 1.

KW Transport; Ion transport; Ionic channel; Voltage-gated channel;  
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;  
 KW Phosphorylation; Multigene family.

FT	DOMAIN	1	186	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	187	208	SEGMENT S1 (POTENTIAL).
FT	TRANSMEM	229	250	SEGMENT S2 (POTENTIAL).
FT	DOMAIN	251	260	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	261	282	SEGMENT S3 (POTENTIAL).
FT	TRANSMEM	295	316	SEGMENT S4 (POTENTIAL).
FT	DOMAIN	317	330	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	331	352	SEGMENT S5 (POTENTIAL).
FT	DOMAIN	365	385	SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT	TRANSMEM	393	414	SEGMENT S6 (POTENTIAL).
FT	DOMAIN	415	858	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	517	520	POLY-SER.
FT	DOMAIN	701	706	POLY-ALA.
SQ	SEQUENCE	858 AA;	95877 MW;	C4B426174ED0DEE4 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 858;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAE 7  
|||||  
Db 348 LVFFAE 353

RESULT 42

KCB1\_PIG

ID KCB1\_PIG STANDARD; PRT; 858 AA.  
AC O18868;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Potassium voltage-gated channel subfamily B member 1 (Potassium  
DE channel Kv2.1) (DRK1).  
GN KCNB1.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lens epithelium;  
RA Rae J.L., Shepard A.R.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: Mediates the voltage-dependent potassium ion  
CC permeability of excitable membranes. Channels open or close in  
CC response to the voltage difference across the membrane, letting K+  
CC ions pass in accordance with their electrochemical gradient (By  
CC similarity).  
CC -!- SUBUNIT: Heteromultimer with KCNG2, KCNG3, KCNG4, KCNS1, KCNS2,  
CC KCNS3 and KCNV2 (By similarity).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is  
CC characterized by a series of positively charged amino acids at  
CC every third position.  
CC -!- DOMAIN: The tail may be important in modulation of channel  
CC activity and/or targeting of the channel to specific subcellular  
CC compartments.  
CC -!- PTM: Phosphorylated on serine residues in the C-terminal  
CC cytoplasmic tail (By similarity).  
CC -!- SIMILARITY: Belongs to the potassium channel family. B (Shab)  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF026006; AAB88809.1; -.

```

DR  HSSP; Q54397; 1BL8.
DR  InterPro; IPR000210; BTB_POZ.
DR  InterPro; IPR005821; Ion_trans.
DR  InterPro; IPR001622; K+channel_pore.
DR  InterPro; IPR003091; K_channel.
DR  InterPro; IPR003131; K_tetra.
DR  InterPro; IPR004350; Kv21channel.
DR  InterPro; IPR003968; Kv_channel.
DR  InterPro; IPR005820; M+channel_nlg.
DR  InterPro; IPR003973; Shab_channel.
DR  Pfam; PF00520; ion_trans; 1.
DR  Pfam; PF02214; K_tetra; 1.
DR  Pfam; PF03521; Kv2channel; 1.
DR  PRINTS; PR00169; KCHANNEL.
DR  PRINTS; PR01514; KV21CHANNEL.
DR  PRINTS; PR01491; KVCHANNEL.
DR  PRINTS; PR01495; SHABCHANNEL.
DR  SMART; SM00225; BTB; 1.
KW  Transport; Ion transport; Ionic channel; Voltage-gated channel;
KW  Potassium channel; Potassium; Potassium transport; Transmembrane;
KW  Phosphorylation; Multigene family.
FT  DOMAIN      1      185      CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM    186     206      SEGMENT S1 (POTENTIAL).
FT  TRANSMEM    229     249      SEGMENT S2 (POTENTIAL).
FT  DOMAIN      250     259      CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM    260     280      SEGMENT S3 (POTENTIAL).
FT  TRANSMEM    295     315      SEGMENT S4 (POTENTIAL).
FT  DOMAIN      316     330      CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM    331     351      SEGMENT S5 (POTENTIAL).
FT  DOMAIN      365     385      SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT  TRANSMEM    393     413      SEGMENT S6 (POTENTIAL).
FT  DOMAIN      414     858      CYTOPLASMIC (POTENTIAL).
FT  DOMAIN      517     520      POLY-SER.
FT  DOMAIN      701     706      POLY-ALA.
SQ  SEQUENCE    858 AA;  96117 MW;  A9E24C3A8E13B491 CRC64;

Query Match          72.5%; Score 29; DB 1; Length 858;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY          2 LVFFAE 7
           |||||
DB         348 LVFFAE 353

```

RESULT 43

KCB1\_RABIT

```

ID  KCB1_RABIT      STANDARD;      PRT;      858 AA.
AC  Q9MZ19;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Potassium voltage-gated channel subfamily B member 1 (Potassium
DE  channel Kv2.1).
GN  KCNB1.
OS  Oryctolagus cuniculus (Rabbit).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=New Zealand white; TISSUE=Corneal endothelium;  
 RA Rae J.L.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Mediates the voltage-dependent potassium ion  
 CC permeability of excitable membranes. Channels open or close in  
 CC response to the voltage difference across the membrane, letting K+  
 CC ions pass in accordance with their electrochemical gradient (By  
 CC similarity).  
 CC -!- SUBUNIT: Heteromultimer with KCNG2, KCNG3, KCNG4, KCNS1, KCNS2,  
 CC KCNS3 and KCNV2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is  
 CC characterized by a series of positively charged amino acids at  
 CC every third position.  
 CC -!- DOMAIN: The tail may be important in modulation of channel  
 CC activity and/or targeting of the channel to specific subcellular  
 CC compartments.  
 CC -!- PTM: Phosphorylated on serine residues in the C-terminal  
 CC cytoplasmic tail (By similarity).  
 CC -!- SIMILARITY: Belongs to the potassium channel family. B (Shab)  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF266507; AAF77058.1; -.  
 DR HSSP; Q54397; 1BL8.  
 DR InterPro; IPR000210; BTB\_POZ.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR001622; K+channel\_pore.  
 DR InterPro; IPR003091; K\_channel.  
 DR InterPro; IPR003131; K\_tetra.  
 DR InterPro; IPR004350; Kv21channel.  
 DR InterPro; IPR003968; Kv\_channel.  
 DR InterPro; IPR005820; M+channel\_nlg.  
 DR InterPro; IPR003973; Shab\_channel.  
 DR Pfam; PF00520; ion\_trans; 1.  
 DR Pfam; PF02214; K\_tetra; 1.  
 DR Pfam; PF03521; Kv2channel; 1.  
 DR PRINTS; PR00169; KCHANNEL.  
 DR PRINTS; PR01514; KV21CHANNEL.  
 DR PRINTS; PR01491; KVCHANNEL.  
 DR PRINTS; PR01495; SHABCHANNEL.  
 DR SMART; SM00225; BTB; 1.  
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;  
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;  
 KW Phosphorylation; Multigene family.  
 FT DOMAIN 1 185 CYTOPLASMIC (POTENTIAL).

FT	TRANSMEM	186	206	SEGMENT S1 (POTENTIAL).
FT	TRANSMEM	229	249	SEGMENT S2 (POTENTIAL).
FT	DOMAIN	250	259	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	260	280	SEGMENT S3 (POTENTIAL).
FT	TRANSMEM	295	315	SEGMENT S4 (POTENTIAL).
FT	DOMAIN	316	330	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	331	351	SEGMENT S5 (POTENTIAL).
FT	DOMAIN	365	385	SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT	TRANSMEM	393	413	SEGMENT S6 (POTENTIAL).
FT	DOMAIN	414	858	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	517	520	POLY-SER.
FT	DOMAIN	701	706	POLY-ALA.
SQ	SEQUENCE	858 AA;	95373 MW;	FDFDCC49B6973C4B CRC64;

Query Match 72.5%; Score 29; DB 1; Length 858;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAE 7  
 |||||  
 Db 348 LVFFAE 353

#### RESULT 44

##### KCB2\_RAT

ID KCB2\_RAT STANDARD; PRT; 907 AA.  
 AC Q63099;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Potassium voltage-gated channel subfamily B member 2 (Potassium channel Kv2.2) (CDRK).  
 GN KCNB2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=92198655; PubMed=1550672;  
 RA Hwang P.M., Glatt C.E., Bredt D.S., Yellen G., Snyder S.H.;  
 RT "A novel K+ channel with unique localizations in mammalian brain:  
 RT molecular cloning and characterization."  
 RL Neuron 8:473-481(1992).  
 RN [2]  
 RP CONCEPTUAL TRANSLATION.  
 RA Sundaram S.;  
 RL Unpublished observations (OCT-2002).  
 CC -!- FUNCTION: Mediates the voltage-dependent potassium ion  
 CC permeability of excitable membranes. Channels open or close in  
 CC response to the voltage difference across the membrane, letting K+  
 CC ions pass in accordance with their electrochemical gradient.  
 CC -!- SUBUNIT: Heteromultimer with KCNS1, KCNS2 and KCNS3 (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: In the brain, the greatest density occurs in

CC the olfactory bulb, followed by the cerebral cortex, hippocampus,  
 CC and cerebellum. In peripheral tissues it is most prominent in  
 CC whole tongue epithelium and circumvallate papillae.  
 CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is  
 CC characterized by a series of positively charged amino acids at  
 CC every third position.  
 CC -!- DOMAIN: The tail may be important in modulation of channel  
 CC activity and/or targeting of the channel to specific subcellular  
 CC compartments.  
 CC -!- PTM: Phosphorylated (Potential).  
 CC -!- SIMILARITY: Belongs to the potassium channel family. B (Shab)  
 CC subfamily.  
 CC -!- CAUTION: This is a conceptual translation, a frameshift was  
 CC introduced in position 755 to restore the similarity with the  
 CC other orthologs.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M77482; AAA40905.1; ALT\_FRAME.

DR HSSP; Q54397; 1BL8.

DR InterPro; IPR000210; BTB\_POZ.

DR InterPro; IPR005821; Ion\_trans.

DR InterPro; IPR001622; K+channel\_pore.

DR InterPro; IPR003091; K\_channel.

DR InterPro; IPR003131; K\_tetra.

DR InterPro; IPR005826; Kv22channel.

DR InterPro; IPR003968; Kv\_channel.

DR InterPro; IPR005820; M+channel\_nlg.

DR InterPro; IPR003973; Shab\_channel.

DR Pfam; PF00520; ion\_trans; 1.

DR Pfam; PF02214; K\_tetra; 1.

DR Pfam; PF03521; Kv2channel; 1.

DR PRINTS; PR00169; KCHANNEL.

DR PRINTS; PR01515; KV22CHANNEL.

DR PRINTS; PR01491; KVCHANNEL.

DR PRINTS; PR01495; SHABCHANNEL.

DR SMART; SM00225; BTB; 1.

KW Transport; Ion transport; Ionic channel; Voltage-gated channel;

KW Potassium channel; Potassium; Potassium transport; Transmembrane;

KW Phosphorylation; Glycoprotein; Multigene family.

FT DOMAIN 1 190 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 191 212 SEGMENT S1 (POTENTIAL).

FT TRANSMEM 233 254 SEGMENT S2 (POTENTIAL).

FT DOMAIN 255 264 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 265 286 SEGMENT S3 (POTENTIAL).

FT TRANSMEM 299 320 SEGMENT S4 (POTENTIAL).

FT DOMAIN 321 334 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 335 356 SEGMENT S5 (POTENTIAL).

FT DOMAIN 369 389 SEGMENT H5 (PORE-FORMING) (POTENTIAL).

FT TRANSMEM 397 418 SEGMENT S6 (POTENTIAL).

FT DOMAIN 419 907 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 907 AA; 102095 MW; B242D9A6753A1295 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 907;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LVFFAE 7  
|||  
Db 352 LVFFAE 357

#### RESULT 45

##### KCB2\_HUMAN

ID KCB2\_HUMAN STANDARD; PRT; 911 AA.  
AC Q92953; Q9BXD3;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Potassium voltage-gated channel subfamily B member 2 (Potassium  
DE channel Kv2.2).  
GN KCNB2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE OF 1-811 FROM N.A.  
RX MEDLINE=98275219; PubMed=9612272;  
RA Schmalz F., Kinsella J., Koh S.D., Vogalis F., Schneider A.,  
RA Flynn E.R., Kenyon J.L., Horowitz B.;  
RT "Molecular identification of a component of delayed rectifier current  
RT in gastrointestinal smooth muscles.";  
RL Am. J. Physiol. 274:G901-G911(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Rae J.L.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: Mediates the voltage-dependent potassium ion  
CC permeability of excitable membranes. Channels open or close in  
CC response to the voltage difference across the membrane, letting K+  
CC ions pass in accordance with their electrochemical gradient.  
CC -!- SUBUNIT: Heteromultimer with KCNS1, KCNS2 and KCNS3 (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is  
CC characterized by a series of positively charged amino acids at  
CC every third position.  
CC -!- DOMAIN: The tail may be important in modulation of channel  
CC activity and/or targeting of the channel to specific subcellular  
CC compartments.  
CC -!- PTM: Phosphorylated (Potential).  
CC -!- SIMILARITY: Belongs to the potassium channel family. B (Shab)  
CC subfamily.  
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a  
CC frameshift in position 768.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U69962; AAB08433.1; ALT\_FRAME.  
 DR EMBL; AF338730; AAK16585.1; -.  
 DR HSSP; Q54397; 1BL8.  
 DR Genew; HGNC:6232; KCNB2.  
 DR MIM; 607738; -.  
 DR GO; GO:0008076; C:voltage-gated potassium channel complex; TAS.  
 DR GO; GO:0005251; F:delayed rectifier potassium channel activity; TAS.  
 DR GO; GO:0006813; P:potassium ion transport; TAS.  
 DR GO; GO:0006940; P:regulation of smooth muscle contraction; TAS.  
 DR InterPro; IPR000210; BTB\_POZ.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR001622; K+channel\_pore.  
 DR InterPro; IPR003091; K\_channel.  
 DR InterPro; IPR003131; K\_tetra.  
 DR InterPro; IPR005826; Kv22channel.  
 DR InterPro; IPR003968; Kv\_channel.  
 DR InterPro; IPR005820; M+channel\_nlg.  
 DR InterPro; IPR003973; Shab\_channel.  
 DR Pfam; PF00520; ion\_trans; 1.  
 DR Pfam; PF02214; K\_tetra; 1.  
 DR Pfam; PF03521; Kv2channel; 1.  
 DR PRINTS; PR00169; KCHANNEL.  
 DR PRINTS; PR01515; KV22CHANNEL.  
 DR PRINTS; PR01491; KVCHANNEL.  
 DR PRINTS; PR01495; SHABCHANNEL.  
 DR SMART; SM00225; BTB; 1.  
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;  
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;  
 KW Phosphorylation; Glycoprotein; Multigene family.  
 FT DOMAIN 1 190 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 191 212 SEGMENT S1 (POTENTIAL).  
 FT TRANSMEM 233 254 SEGMENT S2 (POTENTIAL).  
 FT DOMAIN 255 264 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 265 286 SEGMENT S3 (POTENTIAL).  
 FT TRANSMEM 299 320 SEGMENT S4 (POTENTIAL).  
 FT DOMAIN 321 334 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 335 356 SEGMENT S5 (POTENTIAL).  
 FT DOMAIN 369 389 SEGMENT H5 (PORE-FORMING) (POTENTIAL).  
 FT TRANSMEM 397 418 SEGMENT S6 (POTENTIAL).  
 FT DOMAIN 419 911 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 911 AA; 102562 MW; 2261D7D6280CF81A CRC64;

Query Match 72.5%; Score 29; DB 1; Length 911;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LVFFAE 7  
 |||||

## RESULT 46

## KCB2\_RABIT

ID KCB2\_RABIT STANDARD; PRT; 911 AA.  
AC Q95L11;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Potassium voltage-gated channel subfamily B member 2 (Potassium  
DE channel Kv2.2).  
GN KCNB2.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22286512; PubMed=12399537;  
RA Malysz J., Farrugia G., Ou Y., Szurszewski J.H., Nehra A.,  
RA Gibbons S.J.;  
RT "The Kv2.2 alpha subunit contributes to delayed rectifier K(+)  
RT currents in myocytes from rabbit corpus cavernosum.";  
RL J. Androl. 23:899-910(2002).  
CC -!- FUNCTION: Mediates the voltage-dependent potassium ion  
CC permeability of excitable membranes. Channels open or close in  
CC response to the voltage difference across the membrane, letting K+  
CC ions pass in accordance with their electrochemical gradient (By  
CC similarity).  
CC -!- SUBUNIT: Heteromultimer with KCNS1, KCNS2 and KCNS3 (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is  
CC characterized by a series of positively charged amino acids at  
CC every third position.  
CC -!- DOMAIN: The tail may be important in modulation of channel  
CC activity and/or targeting of the channel to specific subcellular  
CC compartments.  
CC -!- PTM: Phosphorylated (Potential).  
CC -!- SIMILARITY: Belongs to the potassium channel family. B (Shab)  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AY037947; AAK84954.1; -.  
DR InterPro; IPR000210; BTB\_POZ.  
DR InterPro; IPR005821; Ion\_trans.  
DR InterPro; IPR001622; K+channel\_pore.  
DR InterPro; IPR003091; K\_channel.  
DR InterPro; IPR003131; K\_tetra.

DR InterPro; IPR005826; Kv22channel.  
 DR InterPro; IPR003968; Kv\_channel.  
 DR InterPro; IPR005820; M+channel\_nlg.  
 DR InterPro; IPR003973; Shab\_channel.  
 DR Pfam; PF00520; ion\_trans; 1.  
 DR Pfam; PF02214; K\_tetra; 1.  
 DR Pfam; PF03521; Kv2channel; 1.  
 DR PRINTS; PR00169; KCHANNEL.  
 DR PRINTS; PR01515; KV22CHANNEL.  
 DR PRINTS; PR01491; KVCHANNEL.  
 DR PRINTS; PR01495; SHABCHANNEL.  
 DR SMART; SM00225; BTB; 1.  
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;  
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;  
 KW Phosphorylation; Glycoprotein; Multigene family.  
 FT DOMAIN 1 189 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 190 210 SEGMENT S1 (POTENTIAL).  
 FT TRANSMEM 233 253 SEGMENT S2 (POTENTIAL).  
 FT DOMAIN 254 263 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 264 284 SEGMENT S3 (POTENTIAL).  
 FT TRANSMEM 299 320 SEGMENT S4 (POTENTIAL).  
 FT DOMAIN 321 334 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 335 356 SEGMENT S5 (POTENTIAL).  
 FT DOMAIN 369 389 SEGMENT H5 (PORE-FORMING) (POTENTIAL).  
 FT TRANSMEM 397 418 SEGMENT S6 (POTENTIAL).  
 FT DOMAIN 419 911 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 911 AA; 102278 MW; 69353D0664C5D689 CRC64;  
  
 Query Match 72.5%; Score 29; DB 1; Length 911;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LVFFAE 7  
 |||||  
 Db 352 LVFFAE 357

#### RESULT 47

POLG\_ZYMVC

ID POLG\_ZYMVC STANDARD; PRT; 3080 AA.  
 AC P18479; Q89334;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [Contains: N-terminal protein (P1); Helper  
 DE component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa  
 DE protein 1 (6K1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2  
 DE (6K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)  
 DE (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear  
 DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)  
 DE (EC 2.7.7.48); Coat protein (CP)].  
 OS Zucchini yellow mosaic virus (strain California) (ZYMV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
 OC Potyvirus.  
 OX NCBI\_TaxID=117128;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=95146958; PubMed=7844540;  
 RA Wisler G.C., Purcifull D.E., Hiebert E.;  
 RT "Characterization of the P1 protein and coding region of the zucchini  
 RT yellow mosaic virus.";  
 RL J. Gen. Virol. 76:37-45(1995).  
 RN [2]  
 RP SEQUENCE OF 2694-3080 FROM N.A.  
 RX MEDLINE=90236320; PubMed=2185142;  
 RA Gal-On A., Antignus Y., Rosner A., Raccach B.;  
 RT "Nucleotide sequence of the zucchini yellow mosaic virus  
 RT capsid-encoding gene and its expression in Escherichia coli.";  
 RL Gene 87:273-277(1990).  
 CC -!- FUNCTION: Helper component-proteinase is required for aphid  
 CC transmission and also has proteolytic activity.  
 CC -!- FUNCTION: Cytoplasmic inclusion protein has helicase activity. It  
 CC may be involved in replication.  
 CC -!- FUNCTION: Nuclear inclusion protein A has proteolytic activity.  
 CC -!- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is  
 CC further restricted by preferences for the amino acids in P6 - P1'  
 CC that vary with the species of potyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-  
 CC Gln+(Ser or Gly) for the enzyme from tobacco etch virus. The  
 CC natural substrate is the viral polyprotein, but other proteins and  
 CC oligopeptides containing the appropriate consensus sequence are  
 CC also cleaved.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA}(N).  
 CC -!- CATALYTIC ACTIVITY: Hydrolyzes a Gly-|-Gly bond at its own C-  
 CC terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-|-Gly, in the  
 CC processing of the potyviral polyprotein.  
 CC -!- PTM: VPG is covalently linked to the genomic RNA.  
 CC -!- PTM: The viral RNA of potyviruses is expressed as a single  
 CC polyprotein which undergoes posttranslational proteolytic  
 CC processing resulting in the production of at least eight  
 CC individual proteins.  
 CC -!- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.  
 CC -!- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.  
 CC -!- SIMILARITY: Belongs to the potyviruses polyprotein family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L31350; AAA65559.1; -.  
 DR EMBL; M35095; AAA48511.1; -.  
 DR PIR; JH0103; JH0103.  
 DR HSSP; P27958; 1HEI.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR001730; Peptidase\_C4.  
 DR InterPro; IPR001456; Peptidase\_C6.  
 DR InterPro; IPR001592; Poty\_coat.  
 DR InterPro; IPR002540; Poty\_P1.



DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_PSvir.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR Pfam; PF00863; Peptidase\_C4; 1.  
 DR Pfam; PF00851; Peptidase\_C6; 1.  
 DR Pfam; PF00767; Poty\_coat; 1.  
 DR Pfam; PF01577; Poty\_P1; 1.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR PRINTS; PR00966; NIAPOTYPTASE.  
 DR SMART; SM00487; DEXDc; 1.  
 DR SMART; SM00490; HELICc; 1.  
 KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;  
 KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;  
 KW ATP-binding.  
 FT CHAIN 1 304 N-TERMINAL PROTEIN.  
 FT CHAIN 305 766 HELPER COMPONENT PROTEINASE.  
 FT CHAIN 767 ? PROTEIN P3.  
 FT CHAIN ? 1164 6 kDa PROTEIN 1.  
 FT CHAIN 1165 1798 CYTOPLASMIC INCLUSION PROTEIN.  
 FT CHAIN 1799 1851 6 kDa PROTEIN 2.  
 FT CHAIN 1852 2041 GENOME-LINKED PROTEIN.  
 FT CHAIN 2042 2284 NUCLEAR INCLUSION PROTEIN A.  
 FT CHAIN 2285 2801 NUCLEAR INCLUSION PROTEIN B.  
 FT CHAIN 2802 3080 COAT PROTEIN.  
 FT NP\_BIND 1249 1256 ATP (POTENTIAL).  
 FT CONFLICT 2694 2695 LE -> ST (IN REF. 2).  
 FT CONFLICT 2699 2701 IVS -> LFP (IN REF. 2).  
 FT CONFLICT 2811 2811 A -> T (IN REF. 2).  
 FT CONFLICT 2834 2834 G -> S (IN REF. 2).  
 SQ SEQUENCE 3080 AA; 350624 MW; 2A1E501DEA6B9F73 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 3080;  
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFA 6  
 |||||  
 Db 2628 KLVFFA 2633

#### RESULT 48

##### POLG\_ZYMVS

ID POLG\_ZYMVS STANDARD; PRT; 3083 AA.  
 AC O36979;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [Contains: N-terminal protein (P1); Helper  
 DE component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa  
 DE protein 1 (6K1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2  
 DE (6K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)  
 DE (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear  
 DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)  
 DE (EC 2.7.7.48); Coat protein (CP)].  
 OS Zucchini yellow mosaic virus (strain Singapore) (ZYMV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;

OC Potyvirus.  
 OX NCBI\_TaxID=117130;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lee K.C., Wong S.M.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Helper component-proteinase is required for aphid  
 CC transmission and also has proteolytic activity.  
 CC -!- FUNCTION: Cytoplasmic inclusion protein has helicase activity. It  
 CC may be involved in replication.  
 CC -!- FUNCTION: Nuclear inclusion protein A has proteolytic activity.  
 CC -!- CATALYTIC ACTIVITY: Hydrolyzes glutaminy bonds, and activity is  
 CC further restricted by preferences for the amino acids in P6 - P1'  
 CC that vary with the species of potyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-  
 CC Gln+(Ser or Gly) for the enzyme from tobacco etch virus. The  
 CC natural substrate is the viral polyprotein, but other proteins and  
 CC oligopeptides containing the appropriate consensus sequence are  
 CC also cleaved.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA} (N).  
 CC -!- CATALYTIC ACTIVITY: Hydrolyzes a Gly-|-Gly bond at its own C-  
 CC terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-|-Gly, in the  
 CC processing of the potyviral polyprotein.  
 CC -!- PTM: VPG is covalently linked to the genomic RNA.  
 CC -!- PTM: The viral RNA of potyviruses is expressed as a single  
 CC polyprotein which undergoes posttranslational proteolytic  
 CC processing resulting in the production of at least eight  
 CC individual proteins.  
 CC -!- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.  
 CC -!- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.  
 CC -!- SIMILARITY: Belongs to the potyviruses polyprotein family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF014811; AAB72004.2; -.  
 DR HSSP; P27958; 1HEI.  
 DR MEROPS; C04.003; -.  
 DR MEROPS; C06.001; -.  
 DR MEROPS; S30.001; -.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR001730; Peptidase\_C4.  
 DR InterPro; IPR001456; Peptidase\_C6.  
 DR InterPro; IPR001592; Poty\_coat.  
 DR InterPro; IPR002540; Poty\_P1.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_P5vir.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR Pfam; PF00863; Peptidase\_C4; 1.  
 DR Pfam; PF00851; Peptidase\_C6; 1.

DR Pfam; PF00767; Poty\_coat; 1.  
 DR Pfam; PF01577; Poty\_P1; 1.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR PRINTS; PR00966; NIAPOTYPTASE.  
 DR SMART; SM00487; DEXDc; 1.  
 DR SMART; SM00490; HELICc; 1.  
 KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;  
 KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;  
 KW ATP-binding.  
 FT CHAIN 1 307 N-TERMINAL PROTEIN.  
 FT CHAIN 308 769 HELPER COMPONENT PROTEINASE.  
 FT CHAIN 770 1115 PROTEIN P3.  
 FT CHAIN 1116 1167 6 kDa PROTEIN 1.  
 FT CHAIN 1168 1801 CYTOPLASMIC INCLUSION PROTEIN.  
 FT CHAIN 1802 1854 6 kDa PROTEIN 2.  
 FT CHAIN 1855 2044 GENOME-LINKED PROTEIN.  
 FT CHAIN 2045 2287 NUCLEAR INCLUSION PROTEIN A.  
 FT CHAIN 2288 2804 NUCLEAR INCLUSION PROTEIN B.  
 FT CHAIN 2805 3083 COAT PROTEIN.  
 FT NP\_BIND 1252 1259 ATP (POTENTIAL).  
 SQ SEQUENCE 3083 AA; 351028 MW; B4B6C53C14524A88 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 3083;  
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFA 6  
 |||||  
 Db 2631 KLVFFA 2636

#### RESULT 49

##### DYHC\_ANTCR

ID DYHC\_ANTCR STANDARD; PRT; 4466 AA.  
 AC P39057;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Dynein beta chain, ciliary.  
 OS Anthocidaris crassispina (Sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Echinometridae;  
 OC Anthocidaris.  
 OX NCBI\_TaxID=7629;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1192-1204 AND 3324-3340.  
 RC TISSUE=Egg, and Sperm;  
 RX MEDLINE=91326104; PubMed=1830928;  
 RA Ogawa K.;  
 RT "Four ATP-binding sites in the midregion of the beta heavy chain of  
 RT dynein."  
 RL Nature 352:643-645(1991).  
 CC -!- FUNCTION: Force generating protein of eukaryotic cilia and  
 CC flagella. Produces force towards the minus ends of microtubules.  
 CC Dynein has ATPase activity; the force-producing power stroke is  
 CC thought to occur on release of ADP.  
 CC -!- SUBUNIT: Consists of at least two heavy chains (alpha and beta),

CC three intermediate chains and several light chains.

CC -!- SUBCELLULAR LOCATION: Flagellar.

CC -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem  
 CC (which binds cargo and interacts with other dynein components),  
 CC and the head or motor domain. The motor contains six tandemly-  
 CC linked AAA domains in the head, which form a ring. A stalk-like  
 CC structure (formed by two of the coiled coil domains) protrudes  
 CC between AAA 4 and AAA 5 and terminates in a microtubule-binding  
 CC site. A seventh domain may also contribute to this ring; it is not  
 CC clear whether the N-terminus or the C-terminus forms this extra  
 CC domain. There are four well-conserved and two non-conserved ATPase  
 CC sites, one per AAA domain. Probably only one of these (within AAA  
 CC 1) actually hydrolyzes ATP, the others may serve a regulatory  
 CC function.

CC -!- SIMILARITY: Belongs to the dynein heavy chain family.

---

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

---

DR EMBL; D01021; BAA00827.1; -.

DR PIR; S17231; S17231.

DR InterPro; IPR004273; Dynein\_heavy.

DR Pfam; PF03028; Dynein\_heavy; 1.

KW Motor protein; Dynein; Flagellum; Microtubule; ATP-binding; Repeat;

KW Coiled coil.

FT	DOMAIN	1	1813	STEM (BY SIMILARITY).
FT	DOMAIN	1814	2035	AAA 1 (BY SIMILARITY).
FT	DOMAIN	2095	2316	AAA 2 (BY SIMILARITY).
FT	DOMAIN	2422	2669	AAA 3 (BY SIMILARITY).
FT	DOMAIN	2767	3016	AAA 4 (BY SIMILARITY).
FT	DOMAIN	3033	3325	STALK (BY SIMILARITY).
FT	DOMAIN	3409	3636	AAA 5 (BY SIMILARITY).
FT	DOMAIN	3846	4072	AAA 6 (BY SIMILARITY).
FT	DOMAIN	482	502	COILED COIL (POTENTIAL).
FT	DOMAIN	627	643	COILED COIL (POTENTIAL).
FT	DOMAIN	734	805	COILED COIL (POTENTIAL).
FT	DOMAIN	1036	1056	COILED COIL (POTENTIAL).
FT	DOMAIN	1306	1337	COILED COIL (POTENTIAL).
FT	DOMAIN	1443	1468	COILED COIL (POTENTIAL).
FT	DOMAIN	3033	3134	COILED COIL (POTENTIAL).
FT	DOMAIN	3263	3325	COILED COIL (POTENTIAL).
FT	DOMAIN	3573	3642	COILED COIL (POTENTIAL).
FT	NP_BIND	154	161	ATP (POTENTIAL).
FT	NP_BIND	1852	1859	ATP (POTENTIAL).
FT	NP_BIND	2133	2140	ATP (POTENTIAL).
FT	NP_BIND	2460	2467	ATP (POTENTIAL).
FT	NP_BIND	2805	2812	ATP (POTENTIAL).
SQ	SEQUENCE	4466	AA; 511772	MW; C465CC5C6C6D325D CRC64;

Query Match 72.5%; Score 29; DB 1; Length 4466;  
 Best Local Similarity 62.5%; Pred. No. 6.6e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||:| :|  
Db 2520 KLVYFIDD 2527

RESULT 50

DYHC\_TRIGR

ID DYHC\_TRIGR STANDARD; PRT; 4466 AA.  
AC P23098;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Dynein beta chain, ciliary.  
OS Tripneustes gratilla (Hawaian sea urchin).  
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroida; Toxopneustidae;  
OC Tripneustes.  
OX NCBI\_TaxID=7673;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 162-172; 1193-1204; 3240-3259 AND  
RP 3325-3339.  
RC TISSUE=Blastula, and Sperm;  
RX MEDLINE=91326103; PubMed=1830927;  
RA Gibbons I.R., Gibbons B.H., Mocz G., Asai D.J.;  
RT "Multiple nucleotide-binding sites in the sequence of dynein beta  
RT heavy chain.";  
RL Nature 352:640-643(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92020893; PubMed=1833761;  
RA Gibbons I.R., Asai D.J., Ching N.S., Dolecki G.J., Mocz G.,  
RA Phillipson C.A., Ren H., Tang W.Y., Gibbons B.H.;  
RT "A PCR procedure to determine the sequence of large polypeptides by  
RT rapid walking through a cDNA library.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:8563-8567(1991).  
RN [3]  
RP 3D-STRUCTURE MODELING.  
RX MEDLINE=21340771; PubMed=11250194;  
RA Mocz G., Gibbons I.R.;  
RT "Model for the motor component of dynein heavy chain based on homology  
RT to the AAA family of oligomeric ATPases.";  
RL Structure 9:93-103(2001).  
CC -!- FUNCTION: Force generating protein of eukaryotic cilia and  
CC flagella. Produces force towards the minus ends of microtubules.  
CC Dynein has ATPase activity; the force-producing power stroke is  
CC thought to occur on release of ADP.  
CC -!- SUBUNIT: Consists of at least two heavy chains (alpha and beta),  
CC three intermediate chains and several light chains.  
CC -!- SUBCELLULAR LOCATION: Flagellar.  
CC -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem  
CC (which binds cargo and interacts with other dynein components),  
CC and the head or motor domain. The motor contains six tandemly-  
CC linked AAA domains in the head, which form a ring. A stalk-like  
CC structure (formed by two of the coiled coil domains) protrudes  
CC between AAA 4 and AAA 5 and terminates in a microtubule-binding  
CC site. A seventh domain may also contribute to this ring; it is not

```

CC      clear whether the N-terminus or the C-terminus forms this extra
CC      domain. There are four well-conserved and two non-conserved ATPase
CC      sites, one per AAA domain. Probably only one of these (within AAA
CC      1) actually hydrolyzes ATP, the others may serve a regulatory
CC      function.
CC      -!- SIMILARITY: Belongs to the dynein heavy chain family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X59603; CAA42170.1; -.
DR      PIR; S17653; S17653.
DR      PDB; 1HN5; 07-MAR-01.
DR      InterPro; IPR004273; Dynein_heavy.
DR      Pfam; PF03028; Dynein_heavy; 1.
KW      Motor protein; Dynein; Flagellum; Microtubule; ATP-binding; Repeat;
KW      Coiled coil; 3D-structure.
FT      DOMAIN      1      1813      STEM (PROBABLE).
FT      DOMAIN      1814    2035      AAA 1 (PROBABLE).
FT      DOMAIN      2095    2316      AAA 2 (PROBABLE).
FT      DOMAIN      2422    2669      AAA 3 (PROBABLE).
FT      DOMAIN      2767    3016      AAA 4 (PROBABLE).
FT      DOMAIN      3033    3325      STALK (PROBABLE).
FT      DOMAIN      3409    3636      AAA 5 (PROBABLE).
FT      DOMAIN      3846    4072      AAA 6 (PROBABLE).
FT      DOMAIN      733     805      COILED COIL (POTENTIAL).
FT      DOMAIN      1036    1056      COILED COIL (POTENTIAL).
FT      DOMAIN      1306    1337      COILED COIL (POTENTIAL).
FT      DOMAIN      1443    1468      COILED COIL (POTENTIAL).
FT      DOMAIN      3033    3092      COILED COIL (POTENTIAL).
FT      DOMAIN      3263    3325      COILED COIL (POTENTIAL).
FT      DOMAIN      3573    3642      COILED COIL (POTENTIAL).
FT      NP_BIND      154     161      ATP (POTENTIAL).
FT      NP_BIND      1852    1859      ATP (POTENTIAL).
FT      NP_BIND      2133    2140      ATP (POTENTIAL).
FT      NP_BIND      2460    2467      ATP (POTENTIAL).
FT      NP_BIND      2805    2812      ATP (POTENTIAL).
FT      VARIANT      611     615      MISSING.
FT      VARIANT      3357    3357      P -> LLTGNNFFCCFMTA.
SQ      SEQUENCE      4466 AA;  511771 MW;  2A695BF8F336911E CRC64;

```

```

Query Match      72.5%;  Score 29;  DB 1;  Length 4466;
Best Local Similarity  62.5%;  Pred. No. 6.6e+02;
Matches      5;  Conservative      2;  Mismatches      1;  Indels      0;  Gaps      0;

```

```

Qy      1 KLVFFAED 8
        |||:| :|
Db      2520 KLVYFIDD 2527

```

```

RESULT 51
IF1C_TOBAC

```

ID IF1C\_TOBAC STANDARD; PRT; 96 AA.  
 AC P12136;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative translation initiation factor IF-1, chloroplast.  
 GN INFA.  
 OS Nicotiana tabacum (Common tobacco).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamiids; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Bright Yellow 4;  
 RA Shinozaki K., Ohme M., Tanaka M., Wakasugi T., Hayashida N.,  
 RA Matsubayashi T., Zaita N., Chunwongse J., Obokata J.,  
 RA Yamaguchi-Shinozaki K., Ohto C., Torazawa K., Meng B.-Y., Sugita M.,  
 RA Deno H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A.,  
 RA Tohdoh N., Shimada H., Sugiura M.;  
 RT "The complete nucleotide sequence of the tobacco chloroplast genome:  
 RT its gene organization and expression.";  
 RL EMBO J. 5:2043-2049(1986).  
 CC -!- FUNCTION: No specific function has so far been attributed to this  
 CC initiation factor; however, it seems to stimulate more or less all  
 CC the activities of the other two initiation factors, IF-2 and IF-3  
 CC (Potential).  
 CC -!- SUBCELLULAR LOCATION: Chloroplast.  
 CC -!- SIMILARITY: Belongs to the IF-1 family.  
 CC -!- SIMILARITY: Contains 1 S1-like domain.  
 CC -!- CAUTION: This could be encoded by a pseudogene.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z00044; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A05035; A05035.  
 DR HSSP; P02998; 1AH9.  
 DR HAMAP; MF\_00075; atypical; 1.  
 DR InterPro; IPR008994; Nucleic\_acid\_OB.  
 DR InterPro; IPR003029; S1.  
 DR InterPro; IPR006196; S1\_IF1.  
 DR InterPro; IPR004368; TIF\_IF1.  
 DR Pfam; PF00575; S1; 1.  
 DR TIGRFAMs; TIGR00008; infA; 1.  
 DR PROSITE; PS50832; S1\_IF1\_TYPE; 1.  
 KW Initiation factor; Protein biosynthesis; Chloroplast.  
 FT DOMAIN 1 57 S1-LIKE.  
 SQ SEQUENCE 96 AA; 11407 MW; B0BA21A2212EDF6D CRC64;

Query Match 70.0%; Score 28; DB 1; Length 96;

Best Local Similarity 71.4%; Pred. No. 26;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
||:|| |  
Db 84 KLIFFQE 90

RESULT 52

RL7A\_METJA

ID RL7A\_METJA STANDARD; PRT; 117 AA.  
AC P54066;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 50S ribosomal protein L7Ae.  
GN RPL7AE OR MJ1203.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii.";  
RL Science 273:1058-1073(1996).  
CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U67562; AAB99207.1; -.  
DR HSSP; P55769; 1E7K.  
DR TIGR; MJ1203; -.  
DR HAMAP; MF\_00326; -; 1.  
DR InterPro; IPR004038; Ribosomal\_L7A.  
DR InterPro; IPR004037; Ribosomal\_L7Ae.  
DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
DR PRINTS; PR00881; L7ARS6FAMILY.  
DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 117 AA; 12686 MW; 7CAB706F09371BA6 CRC64;



Query Match 70.0%; Score 28; DB 1; Length 117;  
Best Local Similarity 75.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||| |||  
Db 44 KLVIIAED 51

RESULT 53

RL7A\_METAC

ID RL7A\_METAC STANDARD; PRT; 120 AA.  
AC Q8TQL9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 50S ribosomal protein L7Ae.  
GN RPL7AE OR MA1521.  
OS Methanosarcina acetivorans.  
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2214;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
RX MEDLINE=21929760; PubMed=11932238;  
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,  
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,  
RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,  
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,  
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,  
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,  
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,  
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
RA Metcalf W.W., Birren B.;  
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
RT and physiological diversity."  
RL Genome Res. 12:532-542(2002).  
CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE010822; AAM04935.1; -.  
DR HAMAP; MF\_00326; -; 1.  
DR InterPro; IPR004038; Ribosomal\_L7A.  
DR InterPro; IPR004037; Ribosomal\_L7Ae.  
DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
DR PRINTS; PR00881; L7ARS6FAMILY.

DR PROSITE; PS01082; RIBOSOMAL\_L7AE; FALSE\_NEG.  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 120 AA; 12725 MW; C747EB3E29E0F1CC CRC64;

Query Match 70.0%; Score 28; DB 1; Length 120;  
Best Local Similarity 75.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||| |||  
Db 47 KLVLIAED 54

#### RESULT 54

##### RL7A\_METMA

ID RL7A\_METMA STANDARD; PRT; 120 AA.  
AC Q8PU74;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 50S ribosomal protein L7Ae.  
GN RPL7AE OR MM2467.  
OS Methanosarcina mazei (Methanosarcina frisia).  
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2209;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;  
RX MEDLINE=22120827; PubMed=12125824;  
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,  
RA Martinez-Arias R., Henne A., Wiezer A., Baeumer S., Jacobi C.,  
RA Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,  
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
RA Fritz H.-J., Gottschalk G.;  
RT "The genome of Methanosarcina mazei: evidence for lateral gene  
RT transfer between Bacteria and Archaea."  
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; AE013490; AAM32163.1; -.  
DR HAMAP; MF\_00326; -; 1.  
DR InterPro; IPR004038; Ribosomal\_L7A.  
DR InterPro; IPR004037; Ribosomal\_L7Ae.  
DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
DR PRINTS; PR00881; L7ARS6FAMILY.  
DR PROSITE; PS01082; RIBOSOMAL\_L7AE; FALSE\_NEG.  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 120 AA; 12727 MW; 3247EB3E1243C9D6 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 120;  
Best Local Similarity 75.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||| |||  
Db 47 KLVLIAED 54

RESULT 55

RL7A\_PYRHO

ID RL7A\_PYRHO STANDARD; PRT; 123 AA.

AC O59165; Q9V0W4;

DT 15-DEC-1998 (Rel. 37, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE 50S ribosomal protein L7Ae.

GN RPL7AE OR PH1496 OR PYRAB06750 OR PAB0460.

OS Pyrococcus horikoshii, and

OS Pyrococcus abyssi.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI\_TaxID=53953, 29292;

RN [1]

RP SEQUENCE FROM N.A.

RC SPECIES=P.horikoshii; STRAIN=OT3;

RX MEDLINE=98344137; PubMed=9679194;

RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,

RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,

RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,

RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,

RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,

RA Masuchi Y., Shizuya H., Kikuchi H.;

RT "Complete sequence and gene organization of the genome of a hyper-  
thermophilic archaebacterium, Pyrococcus horikoshii OT3.";

RL DNA Res. 5:55-76(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC SPECIES=P.abysii; STRAIN=GE5 / Orsay;

RX MEDLINE=22511545; PubMed=12622808;

RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,

RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,

RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;

RT "An integrated analysis of the genome of the hyperthermophilic  
archaeon Pyrococcus abyssi.";

RL Mol. Microbiol. 47:1495-1512(2003).

CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AP000006; BAA30604.1; ALT\_INIT.  
 DR EMBL; AJ248285; CAB49588.1; ALT\_INIT.  
 DR HSSP; P55769; 1E7K.  
 DR HAMAP; MF\_00326; -; 1.  
 DR InterPro; IPR004038; Ribosomal\_L7A.  
 DR InterPro; IPR004037; Ribosomal\_L7Ae.  
 DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
 DR PRINTS; PR00881; L7ARS6FAMILY.  
 DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 123 AA; 13423 MW; 934817F40CD2D4C5 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 123;  
 Best Local Similarity 75.0%; Pred. No. 33;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||| |||  
 Db 49 KLVIIAED 56

# RESULT 56

## RL7A\_PYRFU

ID RL7A\_PYRFU STANDARD; PRT; 124 AA.  
 AC Q8U160;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L7Ae.  
 GN RPL7AE OR PF1367.  
 OS Pyrococcus furiosus.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=2261;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
 RT "The complete sequence of the Pyrococcus furiosus genome."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE010241; AAL81491.1; -.  
 DR HAMAP; MF\_00326; -; 1.  
 DR InterPro; IPR004038; Ribosomal\_L7A.  
 DR InterPro; IPR004037; Ribosomal\_L7Ae.  
 DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
 DR PRINTS; PR00881; L7ARS6FAMILY.  
 DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.

KW Ribosomal protein; Complete proteome.

SQ SEQUENCE 124 AA; 13526 MW; 932EE7DFEBAFD4C5 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 124;

Best Local Similarity 75.0%; Pred. No. 33;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8

||| |||

Db 50 KLVIIAED 57

# RESULT 57

## RL7A\_SULSO

ID RL7A\_SULSO STANDARD; PRT; 127 AA.

AC P55858;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 50S ribosomal protein L7Ae.

GN RPL7AE OR SSO0091 OR C04\_031.

OS Sulfolobus solfataricus.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI\_TaxID=2287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 35092 / DSM 1617 / P2;

RX MEDLINE=97055432; PubMed=8899719;

RA Sensen C.W., Klenk H.-P., Singh R.K., Allard G., Chan C.C.-Y.,

RA Liu Q.Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,

RA Doolittle W.F., Ragan M.A., Charlebois R.L.;

RT "Organizational characteristics and information content of an

RT archaeal genome: 156 kb of sequence from Sulfolobus solfataricus

RT P2.";

RL Mol. Microbiol. 22:175-191(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 35092 / DSM 1617 / P2;

RX MEDLINE=21332296; PubMed=11427726;

RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,

RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,

RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,

RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,

RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,

RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,

RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;

RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).

CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; Y08257; CAA69560.1; ALT\_INIT.

DR EMBL; AE006649; AAK40449.1; ALT\_INIT.

DR HSSP; P55769; 1E7K.

DR HAMAP; MF\_00326; -; 1.

DR InterPro; IPR004038; Ribosomal\_L7A.

DR InterPro; IPR004037; Ribosomal\_L7Ae.

DR Pfam; PF01248; Ribosomal\_L7Ae; 1.

DR PRINTS; PR00881; L7ARS6FAMILY.

DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.

KW Ribosomal protein; Complete proteome.

SQ SEQUENCE 127 AA; 13736 MW; 3EE30FC81C1A003D CRC64;

Query Match 70.0%; Score 28; DB 1; Length 127;

Best Local Similarity 75.0%; Pred. No. 34;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVFFAED 8

||| |||

Db 49 KLVIIAED 56

#### RESULT 58

##### RL7A\_PYRAE

ID RL7A\_PYRAE STANDARD; PRT; 151 AA.

AC Q8ZTA5;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 50S ribosomal protein L7Ae.

GN RPL7AE OR PAE3347.

OS Pyrobaculum aerophilum.

OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;

OC Thermoproteaceae; Pyrobaculum.

OX NCBI\_TaxID=13773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=IM2 / ATCC 51768 / DSM 7523;

RX MEDLINE=21664397; PubMed=11792869;

RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,

RA Miller J.H.;

RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
aerophilum.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).

CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).

CC

DR EMBL; AE009925; AAL64858.1; -.

DR HAMAP; MF\_00326; -; 1.

DR InterPro; IPR004038; Ribosomal\_L7A.  
 DR InterPro; IPR004037; Ribosomal\_L7Ae.  
 DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
 DR PRINTS; PR00881; L7ARS6FAMILY.  
 DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 151 AA; 16143 MW; AB6FCE4FF27628F9 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 151;  
 Best Local Similarity 75.0%; Pred. No. 40;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||| |||  
 Db 64 KLVLIAED 71

# RESULT 59

## CYSR\_CHICK

ID CYSR\_CHICK STANDARD; PRT; 191 AA.  
 AC P32965;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Cysteine-rich protein 1 (CRP1) (CRP).  
 GN CSRP1 OR CSRP.  
 OS Gallus gallus (Chicken), and  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031, 93934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Chicken; TISSUE=Embryonic fibroblast;  
 RX MEDLINE=94124603; PubMed=8294495;  
 RA Crawford A.W., Pino J.D., Beckerle M.C.;  
 RT "Biochemical and molecular characterization of the chicken  
 RT cysteine-rich protein, a developmentally regulated LIM-domain protein  
 RT that is associated with the actin cytoskeleton."  
 RL J. Cell Biol. 124:117-127(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=C.c.japonica; TISSUE=Embryonic fibroblast;  
 RX MEDLINE=96081967; PubMed=7499425;  
 RA Weiskirchen R., Pino J.D., Macalma T., Bister K., Beckerle M.C.;  
 RT "The cysteine-rich protein family of highly related LIM domain  
 RT proteins."  
 RL J. Biol. Chem. 270:28946-28954(1995).  
 RN [3]  
 RP PARTIAL SEQUENCE.  
 RC SPECIES=Chicken;  
 RX MEDLINE=93107157; PubMed=1469049;  
 RA Sadler I., Crawford A.W., Michelsen J.W., Beckerle M.C.;  
 RT "Zyxin and cCRP: two interactive LIM domain proteins associated with  
 RT the cytoskeleton."  
 RL J. Cell Biol. 119:1573-1587(1992).

RN [4]  
 RP ZINC-BINDING.  
 RX MEDLINE=93281587; PubMed=8506279;  
 RA Michelsen J.W., Schmeichel K.L., Beckerle M.C., Winge D.R.;  
 RT "The LIM motif defines a specific zinc-binding protein domain.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:4404-4408(1993).  
 RN [5]  
 RP MUTAGENESIS.  
 RX MEDLINE=94209279; PubMed=8157637;  
 RA Michelsen J.W., Sewell A.K., Louis H.A., Olsen J.I., Davis D.R.,  
 RA Winge D.R., Beckerle M.C.;  
 RT "Mutational analysis of the metal sites in an LIM domain.";  
 RL J. Biol. Chem. 269:11108-11113(1994).  
 RN [6]  
 RP STRUCTURE BY NMR OF C-TERMINAL LIM DOMAIN.  
 RX MEDLINE=95393167; PubMed=7664053;  
 RA Perez-Alvarado G.C., Miles C., Michelsen J.W., Louis H.A., Winge D.R.;  
 RT "Structure of the carboxy-terminal LIM domain from the cysteine rich  
 RT protein CRP.";  
 RL Nat. Struct. Biol. 1:388-398(1994).  
 CC -!- FUNCTION: Heat stable protein, that interacts with zyxin. May be a  
 CC component of a signal transduction pathway that mediates adhesion-  
 CC stimulated changes in gene expression.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Associates with the actin cytoskeleton.  
 CC -!- TISSUE SPECIFICITY: Most prominent in tissues that are enriched in  
 CC smooth muscle cells, such as gizzard, stomach, and intestine.  
 CC Lower level in the heart, no expression in liver, skeletal muscle,  
 CC or brain.  
 CC -!- DEVELOPMENTAL STAGE: Expression levels increase dramatically  
 CC during smooth muscle maturation.  
 CC -!- DOMAIN: Glycine-rich repeats mediate the association with the  
 CC actin cytoskeleton (Probable).  
 CC -!- SIMILARITY: Contains 2 LIM zinc-binding domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X73831; CAA52053.1; -.  
 DR EMBL; Z28333; CAA82187.1; -.  
 DR PIR; A49648; A49648.  
 DR PIR; B44358; B44358.  
 DR PIR; C44358; C44358.  
 DR PIR; S38879; S38879.  
 DR PDB; 1CTL; 03-JUN-95.  
 DR PDB; 1B8T; 06-MAY-99.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 2.  
 DR ProDom; PD000094; LIM; 2.  
 DR SMART; SM00132; LIM; 2.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 2.  
 DR PROSITE; PS50023; LIM\_DOMAIN\_2; 2.



KW Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc;  
 KW 3D-structure.  
 FT INIT\_MET 0 0  
 FT DOMAIN 9 60 LIM 1.  
 FT DOMAIN 62 77 GLY-RICH.  
 FT DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 117 168 LIM 2.  
 FT DOMAIN 174 185 GLY-RICH.  
 FT STRAND 115 116  
 FT TURN 118 120  
 FT STRAND 123 124  
 FT STRAND 129 132  
 FT TURN 133 134  
 FT STRAND 135 138  
 FT TURN 139 141  
 FT STRAND 143 144  
 FT TURN 145 148  
 FT STRAND 149 150  
 FT STRAND 156 159  
 FT TURN 160 161  
 FT STRAND 162 165  
 FT HELIX 166 171  
 FT TURN 172 172  
 FT TURN 188 191  
 SQ SEQUENCE 191 AA; 20254 MW; BC1A45013C7BDA38 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 191;  
 Best Local Similarity 62.5%; Pred. No. 51;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 | | : | | | :  
 Db 14 KAVYFAEE 21

# RESULT 60

## CYSR\_HUMAN

ID CYSR\_HUMAN STANDARD; PRT; 192 AA.  
 AC P21291;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cysteine-rich protein 1 (CRP1) (CRP).  
 GN CSR1 OR CSR1 OR CYRP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Term placenta;  
 RX MEDLINE=90326508; PubMed=2115670;  
 RA Liehaber S.A., Emery J.G., Urbanek M., Wang X., Cooke N.E.;  
 RT "Characterization of a human cDNA encoding a widely expressed and  
 RT highly conserved cysteine-rich protein with an unusual zinc-finger  
 RT motif."  
 RL Nucleic Acids Res. 18:3871-3879(1990).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92250516; PubMed=1374386;  
 RA Wang X., Lee G., Liebhaber S.A., Cooke N.E.;  
 RT "Human cysteine-rich protein. A member of the LIM/double-finger  
 RT family displaying coordinate serum induction with c-myc.";  
 RL J. Biol. Chem. 267:9176-9184(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Could play a role in neuronal development.  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: Contains 2 LIM zinc-binding domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M33146; AAA35720.1; -.  
 DR EMBL; M76378; AAA58431.1; -.  
 DR EMBL; M76376; AAA58431.1; JOINED.  
 DR EMBL; M76377; AAA58431.1; JOINED.  
 DR EMBL; BC032493; AAH32493.1; -.  
 DR PIR; S12658; S12658.  
 DR HSSP; P32965; 1CTL.  
 DR Genew; HGNC:2469; CSRP1.  
 DR MIM; 123876; -.  
 DR GO; GO:0008270; F:zinc ion binding; TAS.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 2.  
 DR ProDom; PD000094; LIM; 2.

DR SMART; SM00132; LIM; 2.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 2.  
 DR PROSITE; PS50023; LIM\_DOMAIN\_2; 2.  
 KW Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT DOMAIN 9 60 LIM 1.  
 FT DOMAIN 62 77 GLY-RICH.  
 FT DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 118 169 LIM 2.  
 FT DOMAIN 175 186 GLY-RICH.  
 SQ SEQUENCE 192 AA; 20436 MW; 0CC6B9B35DC726FC CRC64;

Query Match 70.0%; Score 28; DB 1; Length 192;  
 Best Local Similarity 62.5%; Pred. No. 51;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 | | : | | | :  
 Db 14 KTVYFAEE 21

# RESULT 61

## CYSR\_RAT

ID CYSR\_RAT STANDARD; PRT; 192 AA.  
 AC P47875;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cysteine-rich protein 1 (CRP1) (CRP).  
 GN CSRP1 OR CSRP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Olfactory bulb;  
 RX MEDLINE=95116342; PubMed=7816640;  
 RA McLaughlin C.R., Tao Q., Abood M.E.;  
 RT "Isolation and developmental expression of a rat cDNA encoding a  
 RT cysteine-rich zinc finger protein."  
 RL Nucleic Acids Res. 22:5477-5483(1994).  
 CC -!- FUNCTION: Could play a role in neuronal development.  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: Contains 2 LIM zinc-binding domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U09567; AAC52157.1; -.  
 DR PIR; S53580; S53580.

DR HSSP; P32965; 1CTL.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 2.  
 DR ProDom; PD000094; LIM; 2.  
 DR SMART; SM00132; LIM; 2.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 2.  
 DR PROSITE; PS50023; LIM\_DOMAIN\_2; 2.  
 KW Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT DOMAIN 9 60 LIM 1.  
 FT DOMAIN 62 77 GLY-RICH.  
 FT DOMAIN 63 68 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 118 169 LIM 2.  
 FT DOMAIN 175 186 GLY-RICH.  
 SQ SEQUENCE 192 AA; 20482 MW; A2EDD41A66B901F2 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 192;  
 Best Local Similarity 62.5%; Pred. No. 51;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 | | : | | | :  
 Db 14 KTVYFAEE 21

# RESULT 62

## RFBF\_SHIFL

ID RFBF\_SHIFL STANDARD; PRT; 296 AA.  
 AC P37782;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE dTDP-rhamnosyl transferase rfbF (EC 2.-.-.-).  
 GN RFBF OR SF2099 OR S2221.  
 OS Shigella flexneri.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Shigella.  
 OX NCBI\_TaxID=623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PE577 / Serotype 2a;  
 RX MEDLINE=94131953; PubMed=7507920;  
 RA Morona R., Mavris M., Fallarino A., Manning P.A.;  
 RT "Characterization of the rfc region of Shigella flexneri."  
 RL J. Bacteriol. 176:733-747(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=301 / Serotype 2a;  
 RX MEDLINE=22272406; PubMed=12384590;  
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
 RA Yu J.;  
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
 RT through comparison with genomes of Escherichia coli K12 and O157."  
 RL Nucleic Acids Res. 30:4432-4441(2002).

RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;  
 RX MEDLINE=22590274; PubMed=12704152;  
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
 RA Schwartz D.C., Blattner F.R.;  
 RT "Complete genome sequence and comparative genomics of Shigella  
 RT flexneri serotype 2a strain 2457T."  
 RL Infect. Immun. 71:2775-2786(2003).  
 CC -!- PATHWAY: Lipopolysaccharide biosynthesis.  
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 2.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X71970; CAA50772.1; -.  
 DR EMBL; AE015225; AAN43638.1; -.  
 DR EMBL; AE016985; AAP17467.1; -.  
 DR PIR; A36966; A36966.  
 DR InterPro; IPR001173; Glyco\_trans\_2.  
 DR InterPro; IPR006446; Rhamnosyltran.  
 DR Pfam; PF00535; Glycos\_transf\_2; 1.  
 DR TIGRFAMs; TIGR01556; rhamnosyltran; 1.  
 KW Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase;  
 KW Complete proteome.  
 FT CONFLICT 203 203 K -> E (IN REF. 1 AND 3).  
 SQ SEQUENCE 296 AA; 33849 MW; E2A75105EF86AE30 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 296;  
 Best Local Similarity 50.0%; Pred. No. 78;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |::|| :|  
 Db 83 KIIFDQD 90

#### RESULT 63

O65A\_DROME

ID O65A\_DROME STANDARD; PRT; 417 AA.  
 AC P82982;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative odorant receptor 65a.  
 GN OR65A.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP CONCEPTUAL TRANSLATION.  
 RA Robertson H.M.;  
 RL Unpublished observations (MAY-2001).  
 CC -!- FUNCTION: Probable role in the odorant response, being an odorant  
 CC receptor.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Belongs to family Dr-or of G-protein coupled  
 CC receptors.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AE003563; -; NOT\_ANNOTATED\_CDS.  
DR FlyBase; FBgn0041625; Or65a.  
DR InterPro; IPR004117; 7tm\_6.  
DR Pfam; PF02949; 7tm\_6; 1.  
KW Hypothetical protein; Transmembrane; G-protein coupled receptor;  
KW Olfaction; Multigene family.  
FT DOMAIN 1 62 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 63 83 1 (POTENTIAL).  
FT DOMAIN 84 98 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 99 119 2 (POTENTIAL).  
FT DOMAIN 120 152 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 153 173 3 (POTENTIAL).  
FT DOMAIN 174 206 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 207 227 4 (POTENTIAL).  
FT DOMAIN 228 290 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 291 311 5 (POTENTIAL).  
FT DOMAIN 312 316 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 317 337 6 (POTENTIAL).  
FT DOMAIN 338 393 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 394 414 7 (POTENTIAL).  
FT DOMAIN 415 417 CYTOPLASMIC (POTENTIAL).  
SQ SEQUENCE 417 AA; 48730 MW; 2112BE9E1E356059 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 417;  
Best Local Similarity 71.4%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
:||||:  
Db 111 RLVFFAQ 117

#### RESULT 64

BCHN\_CHLAU

ID BCHN\_CHLAU STANDARD; PRT; 444 AA.

AC Q9F6X6;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Light-independent protochlorophyllide reductase subunit N

DE (EC 1.18.--.) (LI-POR subunit N) (DPOR subunit N).

GN BCHN.

OS Chloroflexus aurantiacus.

OC Bacteria; Chloroflexi; Chloroflexales; Chloroflexaceae; Chloroflexus.

OX NCBI\_TaxID=1108;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20433268; PubMed=10976061;

RA Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.E.;

RT "Molecular evidence for the early evolution of photosynthesis.";

RL Science 289:1724-1730(2000).

CC -!- FUNCTION: Uses Mg-ATP and reduced ferredoxin to reduce ring D of  
CC protochlorophyllide (Pchlde) to form chlorophyllide a (Chlide).

CC The bchN-bchB pair binds Pchl<sub>ide</sub>. This reaction is light-  
 CC independent.  
 CC -!- PATHWAY: Light-independent bacteriochlorophyll biosynthesis.  
 CC -!- SUBUNIT: Protochlorophyllide reductase is thought to be composed  
 CC of three subunits; bchL, bchN and bchB. Could form a  
 CC heterotetramer of two bchB and two bchN subunits.  
 CC -!- SIMILARITY: Belongs to the bchN / chlN family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF288460; AAG15209.1; -.  
 DR HAMAP; MF\_00352; -; 1.  
 DR InterPro; IPR000510; Oxred\_nitrognse1.  
 DR InterPro; IPR005970; Protochl\_reductN.  
 DR Pfam; PF00148; oxidored\_nitro; 1.  
 DR TIGRFAMs; TIGR01279; DPOR\_bchN; 1.  
 KW Photosynthesis; Bacteriochlorophyll biosynthesis; Oxidoreductase.  
 SQ SEQUENCE 444 AA; 48746 MW; D696C7E3AE735339 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 444;  
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 | :| | | |  
 Db 296 KKIFFASD 303

# RESULT 65

## CRTD\_RHOSH

ID CRTD\_RHOSH STANDARD; PRT; 495 AA.  
 AC Q01671;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Methoxyneurosporene dehydrogenase (EC 1.14.99.-).  
 GN CRTD.  
 OS Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;  
 OC Rhodobacteraceae; Rhodobacter.  
 OX NCBI\_TaxID=1063;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;  
 RX MEDLINE=92307398; PubMed=1612412;  
 RA Gari E., Toledo J.C., Gibert I., Barbe J.;  
 RT "Nucleotide sequence of the methoxyneurosporene dehydrogenase gene  
 RT from Rhodobacter sphaeroides: comparison with other bacterial  
 RT carotenoid dehydrogenases.";  
 RL FEMS Microbiol. Lett. 72:103-108(1992).  
 RN [2]



RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;  
 RX MEDLINE=95238278; PubMed=7721699;  
 RA Lang H.P., Cogdell R.J., Takaichi S., Hunter C.N.;  
 RT "Complete DNA sequence, specific Tn5 insertion map, and gene  
 RT assignment of the carotenoid biosynthesis pathway of Rhodobacter  
 RT sphaeroides.";  
 RL J. Bacteriol. 177:2064-2073(1995).  
 CC -!- FUNCTION: CONVERTS HYDROXYNEUROSPORENE TO DEMETHYLSPHEROIDENE OR  
 CC METHOXYNEUROSPORENE TO SPHEROIDENE.  
 CC -!- COFACTOR: FAD (Probable).  
 CC -!- PATHWAY: Carotenoid and chlorophyll biosynthesis.  
 CC -!- SIMILARITY: Belongs to the phytoene dehydrogenase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X63204; CAA44886.1; -.  
 DR EMBL; AJ010302; CAB38743.1; ALT\_SEQ.  
 DR PIR; S23633; S23633.  
 DR InterPro; IPR002937; Amino\_oxidase.  
 DR InterPro; IPR008150; Bac\_phytoene\_dh.  
 DR InterPro; IPR000205; NAD\_BS.  
 DR Pfam; PF01593; Amino\_oxidase; 1.  
 DR PROSITE; PS00982; PHYTOENE\_DH; 1.  
 KW Photosynthesis; Chlorophyll biosynthesis; Carotenoid biosynthesis;  
 KW Oxidoreductase; FAD; Flavoprotein; NAD.  
 FT NP\_BIND 9 42 FAD (ADP PART) (POTENTIAL).  
 FT CONFLICT 115 115 A -> G (IN REF. 2).  
 FT CONFLICT 157 157 T -> P (IN REF. 2).  
 FT CONFLICT 170 170 L -> M (IN REF. 2).  
 FT CONFLICT 273 273 L -> I (IN REF. 2).  
 FT CONFLICT 440 443 PHGA -> ATGP (IN REF. 1).  
 SQ SEQUENCE 495 AA; 52900 MW; 62227931415B253E CRC64;

Query Match 70.0%; Score 28; DB 1; Length 495;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
 ||||:|  
 Db 331 VFFADD 336

RESULT 66  
 CH60\_CLOTM  
 ID CH60\_CLOTM STANDARD; PRT; 540 AA.  
 AC P48212;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 60 kDa chaperonin (Protein Cpn60) (groEL protein) (HSP-60).

GN GROL OR GROEL OR MOPA.  
 OS Clostridium thermocellum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1515;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCIB 10682;  
 RX MEDLINE=97199381; PubMed=9047357;  
 RA Ciruela A., Cross S., Freedman R.B., Hazlewood G.P.;  
 RT "Sequence and transcriptional analysis of groES and groEL genes from  
 RT the thermophilic bacterium Clostridium thermocellum.";  
 RL Gene 186:143-147(1997).  
 RN [2]  
 RP SEQUENCE OF 1-20.  
 RC STRAIN=NCIB 10682;  
 RX MEDLINE=96257758; PubMed=8687408;  
 RA Cross S.J., Ciruela A., Poomputsa K., Romaniec M.P.M., Freedman R.B.;  
 RT "Thermostable chaperonin from Clostridium thermocellum.";  
 RL Biochem. J. 316:615-622(1996).  
 CC -!- FUNCTION: Prevents misfolding and promotes the refolding and  
 CC proper assembly of unfolded polypeptides generated under stress  
 CC conditions (By similarity).  
 CC -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of  
 CC 7 subunits.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z68137; CAA92242.1; -.  
 DR PIR; S68249; S68249.  
 DR HSSP; P06139; 1GRL.  
 DR HAMAP; MF\_00600; -; 1.  
 DR InterPro; IPR001844; Chaprnin\_Cpn60.  
 DR InterPro; IPR002423; Cpn60/TCP-1.  
 DR InterPro; IPR008950; GroEL-ATPase.  
 DR Pfam; PF00118; cpn60\_TCP1; 1.  
 DR PRINTS; PR00298; CHAPERONIN60.  
 DR PRINTS; PR00304; TCOMPLEXTCP1.  
 DR PROSITE; PS00296; CHAPERONINS\_CPN60; 1.  
 KW Chaperone; ATP-binding.  
 FT INIT\_MET 0 0  
 FT CONFLICT 14 15 LE -> ML (IN REF. 2).  
 FT CONFLICT 20 20 Q -> K (IN REF. 2).  
 SQ SEQUENCE 540 AA; 57343 MW; A638C71A2675C596 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 540;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy            1 KLVFFAED 8  
               |||    |||  
 Db           243 KLVIIAED 250

RESULT 67

YABN\_ECOLI

ID    YABN\_ECOLI            STANDARD;            PRT;    551 AA.  
 AC    P33595; P75638;  
 DT    01-FEB-1994 (Rel. 28, Created)  
 DT    10-OCT-2003 (Rel. 42, Last sequence update)  
 DT    10-OCT-2003 (Rel. 42, Last annotation update)  
 DE    Hypothetical protein yabN.  
 GN    YABN OR B0069.  
 OS    Escherichia coli.  
 OC    Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC    Enterobacteriaceae; Escherichia.  
 OX    NCBI\_TaxID=562;  
 RN    [1]  
 RP    SEQUENCE FROM N.A.  
 RC    STRAIN=K12;  
 RX    MEDLINE=92334977; PubMed=1630901;  
 RA    Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,  
 RA    Isono K., Mizobuchi K., Nakata A.;  
 RT    "Systematic sequencing of the Escherichia coli genome: analysis of  
 RT    the 0-2.4 min region.";  
 RL    Nucleic Acids Res. 20:3305-3308(1992).  
 RN    [2]  
 RP    SEQUENCE FROM N.A.  
 RC    STRAIN=K12 / MG1655;  
 RX    MEDLINE=97426617; PubMed=9278503;  
 RA    Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA    Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA    Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA    Mau B., Shao Y.;  
 RT    "The complete genome sequence of Escherichia coli K-12.";  
 RL    Science 277:1453-1474(1997).  
 CC    -!- SIMILARITY: TO E.COLI YBAE.  
 CC    -----  
 CC    This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC    the European Bioinformatics Institute. There are no restrictions on its  
 CC    use by non-profit institutions as long as its content is in no way  
 CC    modified and this statement is not removed. Usage by and for commercial  
 CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC    -----  
 DR    EMBL; D10483; BAB96638.1; -.  
 DR    EMBL; AE000117; AAC73180.1; -.  
 DR    PIR; E64728; E64728.  
 DR    EcoGene; EG12094; yabN.  
 DR    InterPro; IPR000914; SBP\_bac\_5.  
 DR    Pfam; PF00496; SBP\_bac\_5; 1.  
 KW    Hypothetical protein; Complete proteome.  
 FT    CONFLICT    356       356            V -> A (IN REF. 1).  
 FT    CONFLICT    384       384            L -> F (IN REF. 1).  
 FT    CONFLICT    409       409            Y -> C (IN REF. 1).

SQ SEQUENCE 551 AA; 63975 MW; 0F4FAA1AAFC4E731 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 551;  
Best Local Similarity 71.4%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
||:||||:  
Db 365 LVYFAEE 371

RESULT 68

APB3\_HUMAN

ID APB3\_HUMAN STANDARD; PRT; 575 AA.  
AC O96018; O60483;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Amyloid beta A4 precursor protein-binding family A member 3 (Neuron-specific X11L2 protein) (Neuronal Munc18-1-interacting protein 3)  
DE (Mint-3) (Adapter protein X11gamma).  
GN APBA3 OR MINT3 OR X11L2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99160409; PubMed=10049767;  
RA Tanahashi H., Tabira T.;  
RT "X11L2, a new member of X11 protein family interacts with Alzheimer's  
RT beta-amyloid precursor protein.";  
RL Biochem. Biophys. Res. Commun. 255:663-667(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhart-Schultz K., Brower A., Gordon L., Dias J., Ramirez M.,  
RA Stilwagen S., Phan H., Velasco N., Do L., Regala W., Terry A.,  
RA Garnes J., Danganan L., Erler A., Christensen M., Georgescu A.,  
RA Avila J., Liu S., Attix C., Andreise T., Trankheim M.,  
RA Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R.,  
RA Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D.,  
RA Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP SEQUENCE OF 472-575 FROM N.A.  
RX MEDLINE=99075474; PubMed=9860131;  
RA Okamoto M., Suedhof T.C.;  
RT "Mint 3: a ubiquitous mint isoform that does not bind to munc18-1 or  
RT -2.";  
RL Eur. J. Cell Biol. 77:161-165(1998).  
CC -!- FUNCTION: May modulate processing of the beta-amyloid precursor  
CC protein (APP) and hence formation of beta-APP.  
CC -!- SUBUNIT: Binds to the cytoplasmic domain of amyloid protein (APP)  
CC in vivo.  
CC -!- TISSUE SPECIFICITY: Expressed in all the tissues examined with

CC lower levels in brain and testis.

CC -!- DOMAIN: Composed of an N-terminal domain, a middle  
 CC phosphotyrosine-binding domain (PID/PTB) that mediates binding  
 CC with the cytoplasmic domain of the beta-amyloid precursor protein,  
 CC and two C-terminal PDZ domains thought to attach proteins to the  
 CC plasma membrane.

CC -!- SIMILARITY: Contains 2 PDZ/DHR domains.

CC -!- SIMILARITY: Contains 1 PID domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AB021638; BAA74430.1; -.  
 DR EMBL; AC005954; AAC72275.1; -.  
 DR EMBL; AF029110; AAC17979.1; -.  
 DR PIR; JG0181; JG0181.  
 DR HSSP; Q02410; 1AQC.  
 DR Genew; HGNC:580; APBA3.  
 DR MIM; 604262; -.  
 DR GO; GO:0016020; C:membrane; NAS.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR GO; GO:0007242; P:intracellular signaling cascade; NAS.  
 DR InterPro; IPR001478; PDZ.  
 DR InterPro; IPR006020; PTB\_PID.  
 DR Pfam; PF00595; PDZ; 2.  
 DR Pfam; PF00640; PID; 1.  
 DR SMART; SM00228; PDZ; 2.  
 DR SMART; SM00462; PTB; 1.  
 DR PROSITE; PS50106; PDZ; 2.  
 DR PROSITE; PS01179; PID; 1.

KW Protein transport; Repeat; Polymorphism.

FT DOMAIN 165 171 POLY-SER.  
 FT DOMAIN 217 381 PID.  
 FT DOMAIN 394 480 PDZ 1.  
 FT DOMAIN 485 560 PDZ 2.  
 FT VARIANT 527 527 I -> F (in dbSNP:1045236).  
 FT /FTId=VAR\_011822.  
 FT CONFLICT 505 505 I -> IVRPRPLAPGWGGRAALSTAPEQPPPLSRAPLFLPQ  
 FT (IN REF. 3).

SQ SEQUENCE 575 AA; 61454 MW; 3B910CC74C5F3840 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 575;  
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
 ||:||||  
 Db 332 VFYAED 337

RESULT 69  
 HCYE\_EURCA

ID HCYE\_EURCA STANDARD; PRT; 623 AA.  
 AC P02242;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hemocyanin E chain (HcE).  
 GN HCE.  
 OS Eurypelma californica (American tarantula).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
 OC Mygalomorphae; Theraphosidae; Aphonopelma.  
 OX NCBI\_TaxID=29932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90319102; PubMed=2371273;  
 RA Voll W., Voit R.;  
 RT "Characterization of the gene encoding the hemocyanin subunit e from  
 RT the tarantula Eurypelma californicum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5312-5316(1990).  
 RN [2]  
 RP SEQUENCE.  
 RX MEDLINE=84059635; PubMed=6357986;  
 RA Schneider H.-J., Drexel R., Feldmaier G., Linzen B., Lottspeich F.,  
 RA Henschen A.;  
 RT "Hemocyanins in Spiders, XVIII. Complete amino-acid sequence of  
 RT subunit e from Eurypelma californicum hemocyanin.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1357-1381(1983).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91060544; PubMed=2246235;  
 RA Voit R., Feldmaier-Fuchs G.;  
 RT "Arthropod hemocyanins. Molecular cloning and sequencing of cDNAs  
 RT encoding the tarantula hemocyanin subunits a and e.";  
 RL J. Biol. Chem. 265:19447-19452(1990).  
 RN [4]  
 RP SEQUENCE OF 74-599 FROM N.A.  
 RX MEDLINE=86300721; PubMed=3017715;  
 RA Voit R., Schneider H.-J.;  
 RT "Tarantula hemocyanin mRNA. In vitro translation, cDNA cloning and  
 RT nucleotide sequence corresponding to subunit e.";  
 RL Eur. J. Biochem. 159:23-29(1986).  
 CC -!- FUNCTION: Hemocyanins are copper-containing oxygen carriers  
 CC occurring freely dissolved in the hemolymph of many mollusks and  
 CC arthropods.  
 CC -!- SUBUNIT: Tarantula hemocyanin is a 24-chain polymer with seven  
 CC different chains identified.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- TISSUE SPECIFICITY: Hemolymph.  
 CC -!- MISCELLANEOUS: The two copper ions bound each have 3 nitrogen  
 CC ligands (presumably contributed by His residues) and share a  
 CC bridging ligand (possibly contributed by a Tyr residue) in  
 CC addition to binding oxygen.  
 CC -!- SIMILARITY: Belongs to the tyrosinase family. Hemocyanin  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X16894; CAA34772.1; -.  
 DR EMBL; X04291; CAA27838.1; -.  
 DR EMBL; X16650; CAA34643.1; -.  
 DR EMBL; X16651; CAA34643.1; JOINED.  
 DR EMBL; X16652; CAA34643.1; JOINED.  
 DR EMBL; X16653; CAA34643.1; JOINED.  
 DR EMBL; X16654; CAA34643.1; JOINED.  
 DR EMBL; X16655; CAA34643.1; JOINED.  
 DR EMBL; X16656; CAA34643.1; JOINED.  
 DR EMBL; X16657; CAA34643.1; JOINED.  
 DR PIR; S06701; BHTLE.  
 DR HSSP; P04253; 1OXY.  
 DR InterPro; IPR008922; Di-copper\_centre.  
 DR InterPro; IPR000896; Hemocyanin.  
 DR InterPro; IPR005203; hemocyanin\_C.  
 DR InterPro; IPR005204; hemocyanin\_N.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR002227; Tyrosinase.  
 DR Pfam; PF00372; hemocyanin; 1.  
 DR Pfam; PF03723; hemocyanin\_C; 1.  
 DR Pfam; PF03722; hemocyanin\_N; 1.  
 DR PRINTS; PR00187; HAEMOCYANIN.  
 DR PROSITE; PS00209; HEMOCYANIN\_1; 1.  
 DR PROSITE; PS00210; HEMOCYANIN\_2; 1.  
 DR PROSITE; PS00498; TYROSINASE\_2; 1.  
 KW Transport; Oxygen transport; Copper; Glycoprotein; Hemolymph.  
 FT INIT\_MET 0 0  
 FT METAL 168 168 COPPER 1 (PROBABLE).  
 FT METAL 172 172 COPPER 1 (PROBABLE).  
 FT METAL 199 199 COPPER 1 (PROBABLE).  
 FT METAL 319 319 COPPER 2 (PROBABLE).  
 FT METAL 323 323 COPPER 2 (PROBABLE).  
 FT METAL 359 359 COPPER 2 (PROBABLE).  
 FT DISULFID 528 576 BY SIMILARITY.  
 FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT CONFLICT 29 29 D -> R (IN REF. 2).  
 FT CONFLICT 46 46 C -> D (IN REF. 2).  
 FT CONFLICT 79 79 R -> K (IN REF. 2).  
 FT CONFLICT 90 90 A -> H (IN REF. 2).  
 FT CONFLICT 110 110 MISSING (IN REF. 2).  
 FT CONFLICT 159 159 K -> R (IN REF. 1 AND 3).  
 FT CONFLICT 216 216 R -> H (IN REF. 2).  
 FT CONFLICT 254 254 H -> M (IN REF. 2).  
 FT CONFLICT 307 307 F -> H (IN REF. 2).  
 FT CONFLICT 325 325 M -> MK (IN REF. 2).  
 FT CONFLICT 528 530 CGW -> DGK (IN REF. 2).  
 FT CONFLICT 561 563 NGH -> D (IN REF. 2).  
 FT CONFLICT 602 602 F -> L (IN REF. 3).  
 FT CONFLICT 608 608 V -> P (IN REF. 3).  
 SQ SEQUENCE 623 AA; 71545 MW; E411052A80814004 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 623;

Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|| :| ||  
Db 154 KLAYFKED 161

RESULT 70

HCYF\_EURCA

ID HCYF\_EURCA STANDARD; PRT; 628 AA.  
AC Q9NFL5;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Hemocyanin F chain (HcF).  
GN HCF.  
OS Eurypelma californica (American tarantula).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
OC Mygalomorphae; Theraphosidae; Aphonopelma.  
OX NCBI\_TaxID=29932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Heart;  
RX MEDLINE=20564303; PubMed=10961996;  
RA Voit R., Feldmaier-Fuchs G., Schweikardt T., Decker H., Burmester T.;  
RT "Complete sequence of the 24-mer hemocyanin of the tarantula Eurypelma  
RT californicum. Structure and intramolecular evolution of the  
RT subunits.";  
RL J. Biol. Chem. 275:39339-39344(2000).  
CC -!- FUNCTION: Hemocyanins are copper-containing oxygen carriers  
CC occurring freely dissolved in the hemolymph of many mollusks and  
CC arthropods.  
CC -!- SUBUNIT: Tarantula hemocyanin is a 24-chain polymer with seven  
CC different chains identified.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: Hemolymph.  
CC -!- MISCELLANEOUS: The two copper ions bound each have 3 nitrogen  
CC ligands (presumably contributed by His residues) and share a  
CC bridging ligand (possibly contributed by a Tyr residue) in  
CC addition to binding oxygen.  
CC -!- SIMILARITY: Belongs to the tyrosinase family. Hemocyanin  
CC subfamily.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AJ277491; CAB89496.1; -.  
DR HSSP; P04253; 1LLA.  
DR InterPro; IPR008922; Di-copper\_centre.  
DR InterPro; IPR000896; Hemocyanin.  
DR InterPro; IPR005203; hemocyanin\_C.



DR InterPro; IPR005204; hemocyanin\_N.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR002227; Tyrosinase.  
 DR Pfam; PF00372; hemocyanin; 1.  
 DR Pfam; PF03723; hemocyanin\_C; 1.  
 DR Pfam; PF03722; hemocyanin\_N; 1.  
 DR PRINTS; PR00187; HAEMOCYANIN.  
 DR PROSITE; PS00209; HEMOCYANIN\_1; 1.  
 DR PROSITE; PS00210; HEMOCYANIN\_2; 1.  
 DR PROSITE; PS00498; TYROSINASE\_2; 1.  
 KW Transport; Oxygen transport; Copper; Glycoprotein; Hemolymph.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT DISULFID 533 581 BY SIMILARITY.  
 FT METAL 171 171 COPPER 1 (BY SIMILARITY).  
 FT METAL 175 175 COPPER 1 (BY SIMILARITY).  
 FT METAL 202 202 COPPER 1 (BY SIMILARITY).  
 FT METAL 323 323 COPPER 2 (BY SIMILARITY).  
 FT METAL 327 327 COPPER 2 (BY SIMILARITY).  
 FT METAL 363 363 COPPER 2 (BY SIMILARITY).  
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 628 AA; 72017 MW; 539C94849CC8D4F4 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 628;  
 Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 || :| ||  
 Db 157 KLAYFRED 164

# RESULT 71

## HCYG\_EURCA

ID HCYG\_EURCA STANDARD; PRT; 628 AA.  
 AC Q9NFL4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hemocyanin G chain (HcG).  
 GN HCG.  
 OS Eurypelma californica (American tarantula).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
 OC Mygalomorphae; Theraphosidae; Aphonopelma.  
 OX NCBI\_TaxID=29932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=20564303; PubMed=10961996;  
 RA Voit R., Feldmaier-Fuchs G., Schweikardt T., Decker H., Burmester T.;  
 RT "Complete sequence of the 24-mer hemocyanin of the tarantula Eurypelma  
 RT californicum. Structure and intramolecular evolution of the  
 RT subunits."  
 RL J. Biol. Chem. 275:39339-39344(2000).  
 CC -!- FUNCTION: Hemocyanins are copper-containing oxygen carriers

CC occurring freely dissolved in the hemolymph of many mollusks and  
 CC arthropods.  
 CC -!- SUBUNIT: Tarantula hemocyanin is a 24-chain polymer with seven  
 CC different chains identified.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- TISSUE SPECIFICITY: Hemolymph.  
 CC -!- MISCELLANEOUS: The two copper ions bound each have 3 nitrogen  
 CC ligands (presumably contributed by His residues) and share a  
 CC bridging ligand (possibly contributed by a Tyr residue) in  
 CC addition to binding oxygen.  
 CC -!- SIMILARITY: Belongs to the tyrosinase family. Hemocyanin  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ277492; CAB89497.1; -.  
 DR HSSP; P04253; 10XY.  
 DR InterPro; IPR008922; Di-copper\_centre.  
 DR InterPro; IPR000896; Hemocyanin.  
 DR InterPro; IPR005203; hemocyanin\_C.  
 DR InterPro; IPR005204; hemocyanin\_N.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR002227; Tyrosinase.  
 DR Pfam; PF00372; hemocyanin; 1.  
 DR Pfam; PF03723; hemocyanin\_C; 1.  
 DR Pfam; PF03722; hemocyanin\_N; 1.  
 DR PRINTS; PR00187; HAEMOCYANIN.  
 DR PROSITE; PS00209; HEMOCYANIN\_1; 1.  
 DR PROSITE; PS00210; HEMOCYANIN\_2; 1.  
 DR PROSITE; PS00498; TYROSINASE\_2; 1.  
 KW Transport; Oxygen transport; Copper; Glycoprotein; Hemolymph.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT DISULFID 533 581 BY SIMILARITY.  
 FT METAL 170 170 COPPER 1 (BY SIMILARITY).  
 FT METAL 174 174 COPPER 1 (BY SIMILARITY).  
 FT METAL 201 201 COPPER 1 (BY SIMILARITY).  
 FT METAL 321 321 COPPER 2 (BY SIMILARITY).  
 FT METAL 325 325 COPPER 2 (BY SIMILARITY).  
 FT METAL 361 361 COPPER 2 (BY SIMILARITY).  
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 628 AA; 71679 MW; D63E3384DEA6E7B3 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 628;  
 Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 || :| ||  
 Db 156 KLAYFRED 163

RESULT 72

HCYA\_EURCA

ID HCYA\_EURCA STANDARD; PRT; 630 AA.

AC P14750;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Hemocyanin A chain (HcA).

GN HCA.

OS Eurypelma californica (American tarantula).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;

OC Mygalomorphae; Theraphosidae; Aphonopelma.

OX NCBI\_TaxID=29932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91060544; PubMed=2246235;

RA Voit R., Feldmaier-Fuchs G.;

RT "Arthropod hemocyanins. Molecular cloning and sequencing of cDNAs

RT encoding the tarantula hemocyanin subunits a and e.";

RL J. Biol. Chem. 265:19447-19452(1990).

RN [2]

RP SEQUENCE.

RX MEDLINE=91025623; PubMed=2222854;

RA Schartau W., Metzger W., Sonner P., Geisert H., Storz H.;

RT "Hemocyanins in spiders, XXIII. Complete amino-acid sequence of

RT subunit a of Eurypelma californicum hemocyanin.";

RL Biol. Chem. Hoppe-Seyler 371:557-565(1990).

CC -!- FUNCTION: Hemocyanins are copper-containing oxygen carriers  
 CC occurring freely dissolved in the hemolymph of many mollusks and  
 CC arthropods.

CC -!- SUBUNIT: Tarantula hemocyanin is a 24-chain polymer with seven  
 CC different chains identified.

CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- TISSUE SPECIFICITY: Hemolymph.

CC -!- MISCELLANEOUS: The two copper ions bound each have 3 nitrogen  
 CC ligands (presumably contributed by His residues) and share a  
 CC bridging ligand (possibly contributed by a Tyr residue) in  
 CC addition to binding oxygen.

CC -!- SIMILARITY: Belongs to the tyrosinase family. Hemocyanin  
 CC subfamily.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X16893; CAA34771.1; -.

DR HSSP; P04253; 1LLA.

DR InterPro; IPR008922; Di-copper\_centre.

DR InterPro; IPR000896; Hemocyanin.

DR InterPro; IPR005203; hemocyanin\_C.

DR InterPro; IPR005204; hemocyanin\_N.

DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR002227; Tyrosinase.  
 DR Pfam; PF00372; hemocyanin; 1.  
 DR Pfam; PF03723; hemocyanin\_C; 1.  
 DR Pfam; PF03722; hemocyanin\_N; 1.  
 DR PRINTS; PR00187; HAEMOCYANIN.  
 DR PROSITE; PS00209; HEMOCYANIN\_1; 1.  
 DR PROSITE; PS00210; HEMOCYANIN\_2; 1.  
 DR PROSITE; PS00498; TYROSINASE\_2; 1.  
 KW Transport; Oxygen transport; Copper; Glycoprotein; Hemolymph.  
 FT INIT\_MET 0 0  
 FT METAL 174 174 COPPER 1 (PROBABLE).  
 FT METAL 178 178 COPPER 1 (PROBABLE).  
 FT METAL 205 205 COPPER 1 (PROBABLE).  
 FT METAL 325 325 COPPER 2 (PROBABLE).  
 FT METAL 329 329 COPPER 2 (PROBABLE).  
 FT METAL 365 365 COPPER 2 (PROBABLE).  
 FT DISULFID 537 585 BY SIMILARITY.  
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 630 AA; 72187 MW; 4F6989963C1CC793 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 630;  
 Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 || :| ||  
 Db 160 KLAYFRED 167

# RESULT 73

## MDL1\_CANAL

ID MDL1\_CANAL STANDARD; PRT; 685 AA.  
 AC P97998;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE ATP-dependent permease MDL1.  
 GN MDL1.  
 OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5476;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=491A;  
 RA McCreath K.J.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family. MDR subfamily.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; Y12327; CAA72996.1; -.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR001140; ABC\_TM\_transpt.  
DR InterPro; IPR003439; ABC\_transporter.  
DR Pfam; PF00664; ABC\_membrane; 1.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR ProDom; PD000006; ABC\_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS50929; ABC\_TM1F; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
KW ATP-binding; Transmembrane; Glycoprotein; Transport.  
FT TRANSMEM 96 116 POTENTIAL.  
FT TRANSMEM 149 169 POTENTIAL.  
FT TRANSMEM 256 276 POTENTIAL.  
FT TRANSMEM 354 374 POTENTIAL.  
FT TRANSMEM 381 401 POTENTIAL.  
FT NP\_BIND 475 482 ATP (POTENTIAL).  
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 685 AA; 75868 MW; 46239E214CE1267A CRC64;

Query Match 70.0%; Score 28; DB 1; Length 685;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFA 6  
||:|  
Db 95 KLIFFA 100

#### RESULT 74

##### VIV\_ORYSA

ID VIV\_ORYSA STANDARD; PRT; 728 AA.  
AC P37398;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Viviparous protein homolog.  
GN VP1.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94250843; PubMed=8193305;  
RA Hattori T., Terada T., Hamasuna S.;

RT "Sequence and functional analyses of the rice gene homologous to the  
RT maize Vpl.";  
RL Plant Mol. Biol. 24:805-810(1994).  
CC -!- FUNCTION: Could participate in abscisic acid-regulated gene  
CC expression during seed development.  
CC -!- SIMILARITY: Contains 1 TF-B3 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D16640; BAA04066.1; -.  
DR TRANSFAC; T04785; -.  
DR Gramene; P37398; -.  
DR InterPro; IPR003340; TF\_B3.  
DR Pfam; PF02362; B3; 1.  
KW Developmental protein; Transcription regulation; DNA-binding;  
KW Activator.  
FT DOMAIN 500 685 TF-B3.  
SQ SEQUENCE 728 AA; 75982 MW; F66882079F5FE428 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 728;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
| | | | : |  
Db 46 VFFADD 51

#### RESULT 75

CHLD\_PEA  
ID CHLD\_PEA STANDARD; PRT; 754 AA.  
AC O22437;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Magnesium-chelatase subunit chlD, chloroplast precursor (Mg-  
DE protoporphyrin IX chelatase) (Mg-chelatase subunit D).  
GN CHLD.  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
OX NCBI\_TaxID=3888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Spring; TISSUE=Leaf;  
RA Luo M., Weinstein J.D.;  
RT "Cloning and sequencing of a cDNA encoding the putative Mg-chelatase  
RT subunit D from pea (Pisum sativum L. cv. Spring).";  
RL (In) Plant Gene Register PGR97-139.  
CC -!- FUNCTION: Involved in chlorophyll biosynthesis; introduces a

CC magnesium ion into protoporphyrin IX to yield Mg-protoporphyrin  
CC IX.

CC -!- PATHWAY: Chlorophyll biosynthesis.

CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.

CC -!- SIMILARITY: BELONGS TO THE MG-CHELATASE SUBUNITS D/I FAMILY.

CC -!- SIMILARITY: Contains 1 VWFA domain.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF014399; AAB72194.1; -.

DR PIR; T06249; T06249.

DR InterPro; IPR003593; AAA\_ATPase.

DR InterPro; IPR000523; Mg\_chelatase\_chII.

DR InterPro; IPR002035; VWF\_A.

DR Pfam; PF01078; Mg\_chelatase; 1.

DR SMART; SM00382; AAA; 1.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS50234; VWFA; 1.

KW Photosynthesis; Chlorophyll biosynthesis; ATP-binding; Chloroplast;

KW Transit peptide.

FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).

FT CHAIN ? 754 MAGNESIUM-CHELATASE SUBUNIT CHLD.

FT DOMAIN 551 751 VWFA.

FT DOMAIN 398 449 GLU/PRO-RICH.

FT DOMAIN 405 412 POLY-PRO.

FT DOMAIN 423 432 POLY-GLU.

SQ SEQUENCE 754 AA; 82865 MW; 1D56BA3139080A90 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 754;

Best Local Similarity 71.4%; Pred. No. 2e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAE 7

||:||||:

Db 462 KLLFFAQ 468

#### RESULT 76

##### CHLD\_TOBAC

ID CHLD\_TOBAC STANDARD; PRT; 758 AA.

AC O24133;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Magnesium-chelatase subunit chld, chloroplast precursor (Mg-

DE protoporphyrin IX chelatase) (Mg-chelatase subunit D).

GN CHLD.

OS Nicotiana tabacum (Common tobacco).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

OC lamiids; Solanales; Solanaceae; Nicotiana.

OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98079235; PubMed=9418040;  
 RA Papenbrock J., Graefe S., Kruse E., Haenel F., Grimm B.;  
 RT "Mg-chelatase of tobacco: identification of a Chl D cDNA sequence  
 RT encoding a third subunit, analysis of the interaction of the three  
 RT subunits with the yeast two-hybrid system, and reconstitution of the  
 RT enzyme activity by co-expression of recombinant CHL D, CHL H and CHL  
 RT I.";  
 RL Plant J. 12:981-990(1997).  
 CC -!- FUNCTION: Involved in chlorophyll biosynthesis; introduces a  
 CC magnesium ion into protoporphyrin IX to yield Mg-protoporphyrin  
 CC IX.  
 CC -!- PATHWAY: Chlorophyll biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.  
 CC -!- SIMILARITY: BELONGS TO THE MG-CHELATASE SUBUNITS D/I FAMILY.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y10022; CAA71128.1; -.  
 DR PIR; T02925; T02925.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR000523; Mg\_chelatase\_chII.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01078; Mg\_chelatase; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS50234; VWFA; 1.  
 KW Photosynthesis; Chlorophyll biosynthesis; ATP-binding; Chloroplast;  
 KW Transit peptide.  
 FT TRANSIT 1 62 CHLOROPLAST (POTENTIAL).  
 FT CHAIN 63 758 MAGNESIUM-CHELATASE SUBUNIT CHLD.  
 FT DOMAIN 556 752 VWFA.  
 FT DOMAIN 399 445 GLU/PRO-RICH.  
 FT DOMAIN 409 416 POLY-PRO.  
 FT DOMAIN 427 434 POLY-GLU.  
 SQ SEQUENCE 758 AA; 83012 MW; ED531E85D5FAFEF8 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 758;  
 Best Local Similarity 71.4%; Pred. No. 2e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
 ||:|  
 Db 467 KLLFFAQ 473

RESULT 77  
 CHLD\_ARATH



ID CHLD\_ARATH STANDARD; PRT; 759 AA.  
 AC Q9SJE1; Q9SWY5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Magnesium-chelatase subunit chlD, chloroplast precursor (Mg-  
 DE protoporphyrin IX chelatase) (Mg-chelatase subunit D).  
 GN CHLD OR AT1G08520 OR T27G7.20.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,  
 RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:816-820(2000).  
 RN [2]  
 RP SEQUENCE OF 32-759 FROM N.A.  
 RC STRAIN=cv. C24;  
 RA Green J., Jensen P.E., Gibson L.C.D., Hunter C.N.;  
 RT "Characterization of the magnesium protoporphyrin chelatase chlD  
 RT subunit from Arabidopsis thaliana cv. c24.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Involved in chlorophyll biosynthesis; introduces a  
 CC magnesium ion into protoporphyrin IX to yield Mg-protoporphyrin  
 CC IX.  
 CC -!- PATHWAY: Chlorophyll biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.  
 CC -!- SIMILARITY: BELONGS TO THE MG-CHELATASE SUBUNITS D/I FAMILY.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to erroneous  
 CC gene model prediction.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AC006932; AAF22895.1; ALT\_SEQ.  
DR EMBL; AF083555; AAD52031.1; -.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR000523; Mg\_chelatase\_chII.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01078; Mg\_chelatase; 1.  
DR SMART; SM00382; AAA; 1.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS50234; VWFA; 1.  
KW Photosynthesis; Chlorophyll biosynthesis; ATP-binding; Chloroplast;  
KW Transit peptide.  
FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).  
FT CHAIN ? 759 MAGNESIUM-CHELATASE SUBUNIT CHLD.  
FT DOMAIN 557 753 VWFA.  
FT DOMAIN 411 418 POLY-PRO.  
FT DOMAIN 425 438 POLY-GLU.  
FT CONFLICT 70 70 D -> E (IN REF. 2).  
FT CONFLICT 150 150 D -> N (IN REF. 2).  
FT CONFLICT 270 270 G -> S (IN REF. 2).  
FT CONFLICT 285 286 MISSING (IN REF. 2).  
FT CONFLICT 314 314 R -> S (IN REF. 2).  
FT CONFLICT 379 379 E -> K (IN REF. 2).  
FT CONFLICT 426 426 E -> D (IN REF. 2).  
SQ SEQUENCE 759 AA; 83154 MW; A4169319E1247BD1 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 759;  
Best Local Similarity 71.4%; Pred. No. 2e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
||:|:|:  
Db 468 KLLFFAQ 474

#### RESULT 78

##### POLG\_HCVB

ID POLG\_HCVB STANDARD; PRT; 3898 AA.  
AC P21530;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Genome polyprotein.  
OS Hog cholera virus (strain Brescia) (Swine fever virus).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Pestivirus.  
OX NCBI\_TaxID=11098;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90281581; PubMed=2162104;  
RA Moormann R.J.M., Warmerdam P.A.M., van der Meer B., Schaaper W.M.M.,  
RA Wensvoort G., Hulst M.M.;  
RT "Molecular cloning and nucleotide sequence of hog cholera virus  
RT strain Brescia and mapping of the genomic region encoding envelope

```

RT    protein E1.";
RL    Virology 177:184-198(1990).
CC    -!- FUNCTION: PESTIVIRUS P80 (P125) MAY BE A BIFUNCTIONAL PROTEIN
CC        WITH HELICASE AND PROTEASE ACTIVITY.
CC    -!- SUBCELLULAR LOCATION: THE GP51-GP54 PROTEIN IS ANCHORED TO THE
CC        VIRAL ENVELOPE.
CC    -!- SIMILARITY: TO BOVINE VIRAL DIARRHEA VIRUS GENOME POLYPROTEIN.
CC    -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S31.
CC    -----
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    -----
DR    EMBL; M31768; AAA43843.1; -.
DR    PIR; A35317; GNWVHB.
DR    HSSP; P27958; 1A1V.
DR    MEROPS; C53.001; -.
DR    MEROPS; S31.001; -.
DR    InterPro; IPR001410; DEAD.
DR    InterPro; IPR002166; HCV_RdRP.
DR    InterPro; IPR001650; Helicase_C.
DR    InterPro; IPR008751; Peptidase_C53.
DR    InterPro; IPR000280; Peptidase_S31.
DR    InterPro; IPR007095; RNA_pol_DS_PS.
DR    InterPro; IPR007094; RNA_pol_PSvir.
DR    InterPro; IPR001568; RNase_T2.
DR    Pfam; PF00271; helicase_C; 1.
DR    Pfam; PF05550; Peptidase_C53; 1.
DR    Pfam; PF05578; Peptidase_S31; 1.
DR    Pfam; PF00998; Viral_RdRP; 1.
DR    PRINTS; PR00729; CDVENDOPTASE.
DR    SMART; SM00487; DEXDc; 1.
DR    SMART; SM00490; HELICc; 1.
DR    PROSITE; PS00531; RNASE_T2_2; UNKNOWN_1.
KW    Polyprotein; Glycoprotein; Transmembrane; Hydrolase; Serine protease;
KW    Helicase.
FT    CHAIN      ?1      2267      CAPSID PROTEIN C (POTENTIAL).
FT    CHAIN      ?268     2500      GP42 (E2) (POTENTIAL).
FT    CHAIN      ?501     ?689      GP31 (E3) (POTENTIAL).
FT    CHAIN      ?690     ?1060     GP51-GP54 (ENVELOPE PROTEIN E1).
FT    CHAIN      ?1611    ?2111    P80 (POTENTIAL).
FT    TRANSMEM   1032     1048      POTENTIAL.
FT    ACT_SITE   1658     1658      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT    ACT_SITE   1695     1695      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT    ACT_SITE   1752     1752      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT    CARBOHYD   157      157      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    CARBOHYD   269      269      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    CARBOHYD   278      278      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    CARBOHYD   332      332      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    CARBOHYD   362      362      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    CARBOHYD   410      410      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    CARBOHYD   425      425      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    CARBOHYD   500      500      N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT	CARBOHYD	594	594	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	805	805	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	810	810	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	874	874	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	918	918	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	949	949	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1713	1713	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2134	2134	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2217	2217	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2419	2419	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2494	2494	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2787	2787	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2815	2815	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2891	2891	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3103	3103	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3211	3211	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3316	3316	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3689	3689	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3698	3698	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3794	3794	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	3898	AA; 438423	MW; EC6EB207A09D59FD CRC64;

Query Match 70.0%; Score 28; DB 1; Length 3898;  
 Best Local Similarity 62.5%; Pred. No. 9.8e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 | :|| ||  
 Db 1525 KRIFFRED 1532

# RESULT 79

## DYH9\_HUMAN

ID DYH9\_HUMAN STANDARD; PRT; 4486 AA.  
 AC Q9NYC9; O95494; Q9NQ28;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ciliary dynein heavy chain 9 (Axonemal beta dynein heavy chain 9).  
 GN DNAH9 OR DNAH17L OR DNAL1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Reed W., Moats-Staats B.M., Carson J.L., Leigh M.W., Collier A.M.;  
 RT "A ciliary dynein heavy chain whose expression is upregulated in  
 RT differentiating airway epithelium."  
 RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Nasal epithelium;  
 RX MEDLINE=21145582; PubMed=11247663;  
 RA Bartoloni L., Blouin J.L., Maiti A.K., Sainsbury A., Rossier C.,  
 RA Gehrig C., She J.X., Marron M.P., Lander E.S., Meeks M., Chung E.,  
 RA Armengot M., Jorissen M., Scott H.S., Delozier-Blanchet C.D.,

RA Gardiner R.M., Antonarakis S.E.;  
 RT "Axonemal beta heavy chain dynein DNAH9: cDNA sequence, genomic  
 RT structure, and investigation of its role in primary ciliary  
 RT dyskinesia.";  
 RL Genomics 72:21-33(2001).  
 RN [3]  
 RP SEQUENCE OF 1874-1974 FROM N.A.  
 RC TISSUE=Nasal polyps;  
 RA Maiti A.K., Mattei M.-G., Jorissen M., Volz A., Ziegler A.,  
 RA Bouvagnet P.;  
 RT "Chromosomal localization of human dynein heavy chain genes.";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=20558134; PubMed=11104725;  
 RA Reed W., Carson J.L., Moats-Staats B.M., Lucier T., Hu P.C.,  
 RA Brighton L., Gambling T.M., Huang C.H., Leigh M.W., Collier A.M.;  
 RT "Characterization of an axonemal dynein heavy chain expressed early in  
 RT airway epithelial ciliogenesis.";  
 RL Am. J. Respir. Cell Mol. Biol. 23:734-741(2000).  
 CC -!- FUNCTION: Force generating protein of respiratory cilia. Produces  
 CC force towards the minus ends of microtubules. Dynein has ATPase  
 CC activity; the force-producing power stroke is thought to occur on  
 CC release of ADP.  
 CC -!- SUBUNIT: Consists of at least two heavy chains and a number of  
 CC intermediate and light chains.  
 CC -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem  
 CC (which binds cargo and interacts with other dynein components),  
 CC and the head or motor domain. The motor contains six tandemly-  
 CC linked AAA domains in the head, which form a ring. A stalk-like  
 CC structure (formed by two of the coiled coil domains) protrudes  
 CC between AAA 4 and AAA 5 and terminates in a microtubule-binding  
 CC site. A seventh domain may also contribute to this ring; it is not  
 CC clear whether the N-terminus or the C-terminus forms this extra  
 CC domain. There are four well-conserved and two non-conserved ATPase  
 CC sites, one per AAA domain. Probably only one of these (within AAA  
 CC 1) actually hydrolyzes ATP, the others may serve a regulatory  
 CC function.  
 CC -!- SIMILARITY: Belongs to the dynein heavy chain family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF257737; AAF69004.1; -.  
 DR EMBL; AJ404468; CAB94756.1; -.  
 DR EMBL; AJ132088; CAA10561.1; -.  
 DR Genew; HGNC:2953; DNAH9.  
 DR MIM; 603330; -.  
 DR GO; GO:0005858; C:axonemal dynein complex; NAS.  
 DR GO; GO:0005856; C:cytoskeleton; TAS.  
 DR GO; GO:0003775; F:axonemal motor activity; NAS.  
 DR GO; GO:0006928; P:cell motility; TAS.

DR GO; GO:0007283; P:spermatogenesis; TAS.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR004273; Dynein\_heavy.  
 DR Pfam; PF03028; Dynein\_heavy; 1.  
 DR SMART; SM00382; AAA; 3.  
 KW Motor protein; Dynein; Microtubule; ATP-binding; Repeat; Coiled coil.  
 FT DOMAIN 1 1831 STEM (BY SIMILARITY).  
 FT DOMAIN 1832 2053 AAA 1 (BY SIMILARITY).  
 FT DOMAIN 2113 2334 AAA 2 (BY SIMILARITY).  
 FT DOMAIN 2440 2688 AAA 3 (BY SIMILARITY).  
 FT DOMAIN 2787 3036 AAA 4 (BY SIMILARITY).  
 FT DOMAIN 3051 3341 STALK (BY SIMILARITY).  
 FT DOMAIN 3429 3656 AAA 5 (BY SIMILARITY).  
 FT DOMAIN 3866 4092 AAA 6 (BY SIMILARITY).  
 FT DOMAIN 381 410 COILED COIL (POTENTIAL).  
 FT DOMAIN 504 529 COILED COIL (POTENTIAL).  
 FT DOMAIN 639 662 COILED COIL (POTENTIAL).  
 FT DOMAIN 752 823 COILED COIL (POTENTIAL).  
 FT DOMAIN 1326 1355 COILED COIL (POTENTIAL).  
 FT DOMAIN 3051 3154 COILED COIL (POTENTIAL).  
 FT DOMAIN 3285 3341 COILED COIL (POTENTIAL).  
 FT DOMAIN 3640 3675 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1870 1877 ATP (POTENTIAL).  
 FT NP\_BIND 2151 2158 ATP (POTENTIAL).  
 FT NP\_BIND 2478 2485 ATP (POTENTIAL).  
 FT NP\_BIND 2825 2832 ATP (POTENTIAL).  
 FT CONFLICT 2505 2505 L -> V (IN REF. 2).  
 FT CONFLICT 3678 3678 T -> A (IN REF. 2).  
 FT CONFLICT 4374 4374 I -> M (IN REF. 2).  
 SQ SEQUENCE 4486 AA; 511927 MW; 996EDFFDEB0B3EB1 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 4486;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||::| :|  
 Db 2539 KLIYFIDD 2546

RESULT 80  
 DYHB\_HUMAN  
 ID DYHB\_HUMAN STANDARD; PRT; 4523 AA.  
 AC Q96DT5; Q9UJ82;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ciliary dynein heavy chain 11 (Axonemal beta dynein heavy chain 11).  
 GN DNAH11.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., VARIANT PCD GLN-3004, AND VARIANTS LEU-34;  
 RP ARG-639; CYS-654; ALA-1023; THR-1038; GLY-1640; ASN-2641; THR-3474;  
 RP VAL-3715; PRO-3765 AND ILE-4177.

RC TISSUE=Nasal epithelium, and Testis;  
 RX MEDLINE=22155903; PubMed=12142464;  
 RA Bartoloni L., Blouin J.L., Pan Y., Gehrig C., Maiti A.K., Scamuffa N.,  
 RA Rossier C., Jorissen M., Armengot M., Meeks M., Mitchison H.M.,  
 RA Chung E.M., Delozier-Blanchet C.D., Craigen W.J., Antonarakis S.E.;  
 RT "Mutations in the DNAH11 (axonemal heavy chain dynein type 11) gene  
 RT cause one form of situs inversus totalis and most likely primary  
 RT ciliary dyskinesia.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10282-10286(2002).  
 RN [2]  
 RP SEQUENCE OF 1904-2004 FROM N.A.  
 RC TISSUE=Nasal polyps;  
 RA Maiti A.K., Mattei M.-G., Jorissen M., Volz A., Ziegler A.,  
 RA Bouvagnet P.;  
 RT "Chromosomal localization of human dynein heavy chain genes.";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Force generating protein of respiratory cilia. Produces  
 CC force towards the minus ends of microtubules. Dynein has ATPase  
 CC activity; the force-producing power stroke is thought to occur on  
 CC release of ADP.  
 CC -!- SUBUNIT: Consists of at least two heavy chains and a number of  
 CC intermediate and light chains.  
 CC -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem  
 CC (which binds cargo and interacts with other dynein components),  
 CC and the head or motor domain. The motor contains six tandemly-  
 CC linked AAA domains in the head, which form a ring. A stalk-like  
 CC structure (formed by two of the coiled coil domains) protrudes  
 CC between AAA 4 and AAA 5 and terminates in a microtubule-binding  
 CC site. A seventh domain may also contribute to this ring; it is not  
 CC clear whether the N-terminus or the C-terminus forms this extra  
 CC domain. There are four well-conserved and two non-conserved ATPase  
 CC sites, one per AAA domain. Probably only one of these (within AAA  
 CC 1) actually hydrolyzes ATP, the others may serve a regulatory  
 CC function.  
 CC -!- DISEASE: Defects in DNAH11 are a cause of primary ciliary  
 CC dyskinesia (PCD) [MIM:242650]; also known as immotile cilia  
 CC syndrome 1 (ICS1). The phenotype of this autosomal recessive  
 CC disease is characterized by axonemal abnormalities of respiratory  
 CC cilia and sperm tails leading to bronchiectasis and sinusitis,  
 CC which are sometimes associated with situs inversus (Kartagener  
 CC syndrome) and male sterility.  
 CC -!- SIMILARITY: Belongs to the dynein heavy chain family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ320497; CAC60121.1; -.  
 DR EMBL; AJ132087; CAA10560.1; -.  
 DR Genew; HGNC:2942; DNAH11.  
 DR MIM; 603339; -.  
 DR MIM; 242650; -.  
 DR GO; GO:0005858; C:axonemal dynein complex; NAS.

DR GO; GO:0003775; F:axonemal motor activity; NAS.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR004273; Dynein\_heavy.  
 DR Pfam; PF03028; Dynein\_heavy; 1.  
 DR SMART; SM00382; AAA; 4.  
 KW Motor protein; Dynein; Microtubule; ATP-binding; Repeat; Coiled coil;  
 KW Polymorphism; Disease mutation.  
 FT DOMAIN 1 1861 STEM (BY SIMILARITY).  
 FT DOMAIN 1862 2083 AAA 1 (BY SIMILARITY).  
 FT DOMAIN 2143 2373 AAA 2 (BY SIMILARITY).  
 FT DOMAIN 2479 2726 AAA 3 (BY SIMILARITY).  
 FT DOMAIN 2824 3073 AAA 4 (BY SIMILARITY).  
 FT DOMAIN 3079 3410 STALK (BY SIMILARITY).  
 FT DOMAIN 3466 3693 AAA 5 (BY SIMILARITY).  
 FT DOMAIN 3903 4129 AAA 6 (BY SIMILARITY).  
 FT DOMAIN 1274 1327 COILED COIL (POTENTIAL).  
 FT DOMAIN 3079 3143 COILED COIL (POTENTIAL).  
 FT DOMAIN 3319 3410 COILED COIL (POTENTIAL).  
 FT DOMAIN 3675 3710 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1900 1907 ATP (POTENTIAL).  
 FT NP\_BIND 2181 2188 ATP (POTENTIAL).  
 FT NP\_BIND 2517 2524 ATP (POTENTIAL).  
 FT NP\_BIND 2862 2869 ATP (POTENTIAL).  
 FT VARIANT 34 34 E -> L (requires 2 nucleotide  
 substitutions).  
 FT /FTId=VAR\_013851.  
 FT VARIANT 639 639 Q -> R.  
 FT /FTId=VAR\_013852.  
 FT VARIANT 654 654 S -> C.  
 FT /FTId=VAR\_013853.  
 FT VARIANT 1023 1023 V -> A.  
 FT /FTId=VAR\_013854.  
 FT VARIANT 1038 1038 A -> T.  
 FT /FTId=VAR\_013855.  
 FT VARIANT 1640 1640 D -> G.  
 FT /FTId=VAR\_013856.  
 FT VARIANT 2641 2641 S -> N.  
 FT /FTId=VAR\_013857.  
 FT VARIANT 2682 2682 I -> V.  
 FT /FTId=VAR\_013858.  
 FT VARIANT 3004 3004 R -> Q (in PCD; not proven to be  
 pathogenic).  
 FT /FTId=VAR\_013859.  
 FT VARIANT 3474 3474 A -> T.  
 FT /FTId=VAR\_013860.  
 FT VARIANT 3715 3715 L -> V.  
 FT /FTId=VAR\_013861.  
 FT VARIANT 3765 3765 S -> P.  
 FT /FTId=VAR\_013862.  
 FT VARIANT 4177 4177 T -> I.  
 FT /FTId=VAR\_013863.  
 SQ SEQUENCE 4523 AA; 520969 MW; 7C9A71C95B296B89 CRC64;

Query Match 70.0%; Score 28; DB 1; Length 4523;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;



Qy            1 KLVFFAED 8  
             ||::| :|  
Db            2577 KLIYFIDD 2584

RESULT 81

VNS2\_CVBM

ID    VNS2\_CVBM            STANDARD;            PRT;    109 AA.  
AC    P15774;  
DT    01-APR-1990 (Rel. 14, Created)  
DT    01-APR-1990 (Rel. 14, Last sequence update)  
DT    10-OCT-2003 (Rel. 42, Last annotation update)  
DE    Nonstructural protein 2 (Nonstructural 12.7 kDa protein).  
GN    NS2.  
OS    Bovine coronavirus (strain Mebus) (BCoV) (BCV), and  
OS    Bovine coronavirus (strain F15) (BCoV) (BCV).  
OC    Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC    Coronaviridae; Coronavirus.  
OX    NCBI\_TaxID=11132, 11129;  
RN    [1]  
RP    SEQUENCE FROM N.A.  
RC    STRAIN=Mebus;  
RX    MEDLINE=90320120; PubMed=2142556;  
RA    Abraham S., Kienzle T.E., Lapps W.E., Brian D.A.;  
RT    "Sequence and expression analysis of potential nonstructural proteins  
RT    of 4.9, 4.8, 12.7, and 9.5 kDa encoded between the spike and membrane  
RT    protein genes of the bovine coronavirus."  
RL    Virology 177:488-495(1990).  
RN    [2]  
RP    SEQUENCE FROM N.A.  
RC    STRAIN=F15;  
RX    MEDLINE=90206809; PubMed=2320429;  
RA    Woloszyn N., Boireau P., Laporte J.;  
RT    "Nucleotide sequence of the bovine enteric coronavirus BECV F15 mRNA  
RT    5 and mRNA 6 unique regions."  
RL    Nucleic Acids Res. 18:1303-1303(1990).  
CC    -!- SIMILARITY: Belongs to the coronavirus NS2 protein family.  
CC    -----  
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC    the European Bioinformatics Institute. There are no restrictions on its  
CC    use by non-profit institutions as long as its content is in no way  
CC    modified and this statement is not removed. Usage by and for commercial  
CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC    -----  
DR    EMBL; X51347; CAA35740.1; -.  
DR    EMBL; U00735; AAA42913.1; -.  
DR    PIR; S08408; MNIHB2.  
DR    InterPro; IPR006841; Corona\_NS2.  
DR    Pfam; PF04753; Corona\_NS2; 1.  
KW    Nonstructural protein.  
SQ    SEQUENCE    109 AA; 12806 MW; 08B7CA339A1BD051 CRC64;

Query Match            67.5%;    Score 27;    DB 1;    Length 109;  
Best Local Similarity    75.0%;    Pred. No. 50;  
Matches        6;    Conservative    0;    Mismatches        2;    Indels        0;    Gaps        0;

Qy           1 KLVFFAED 8  
              |||| | |  
Db           89 KLVFLAVD 96

RESULT 82

VNS2\_CVHOC

ID VNS2\_CVHOC           STANDARD;           PRT;   109 AA.  
AC Q04853;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Nonstructural protein 2 (Nonstructural 12.9 kDa protein).  
GN NS2.  
OS Human coronavirus (strain OC43) (HCoV-OC43).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC Coronaviridae; Coronavirus.  
OX NCBI\_TaxID=31631;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93297129; PubMed=8517026;  
RA Mounir S., Talbot P.J.;  
RT "Human coronavirus OC43 RNA 4 lacks two open reading frames located  
RT downstream of the S gene of bovine coronavirus.";  
RL Virology 192:355-360(1993).  
CC -!- FUNCTION: May act as membrane-anchoring region for structural  
CC       proteins during virus assembly, or play a role in membrane  
CC       association of the viral polymerase during replication.  
CC -!- SIMILARITY: Belongs to the coronavirus NS2 protein family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M99576; AAA02569.1; -.  
DR PIR; A44275; A44275.  
DR InterPro; IPR006841; Corona\_NS2.  
DR Pfam; PF04753; Corona\_NS2; 1.  
KW Nonstructural protein.  
SQ SEQUENCE   109 AA;   12935 MW;   8F06A6E9E30B1667 CRC64;

Query Match                   67.5%;   Score 27;   DB 1;   Length 109;  
Best Local Similarity   75.0%;   Pred. No. 50;  
Matches   6;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy           1 KLVFFAED 8  
              |||| | |  
Db           89 KLVFLAVD 96

RESULT 83

RL7A\_THEAC

ID RL7A\_THEAC STANDARD; PRT; 121 AA.  
AC Q9HJ56;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 50S ribosomal protein L7Ae.  
GN RPL7AE OR TA1116.  
OS Thermoplasma acidophilum.  
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
OC Thermoplasmataceae; Thermoplasma.  
OX NCBI\_TaxID=2303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 1728;  
RX MEDLINE=20479972; PubMed=11029001;  
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,  
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;  
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma  
RT acidophilum.";  
RL Nature 407:508-513(2000).  
CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AL445066; CAC12243.1; -.  
DR HSSP; P55769; 1E7K.  
DR HAMAP; MF\_00326; -; 1.  
DR InterPro; IPR004038; Ribosomal\_L7A.  
DR InterPro; IPR004037; Ribosomal\_L7Ae.  
DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
DR PRINTS; PR00881; L7ARS6FAMILY.  
DR PROSITE; PS01082; RIBOSOMAL\_L7AE; FALSE\_NEG.  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 121 AA; 13435 MW; 4156F6228E0E8B51 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 121;  
Best Local Similarity 75.0%; Pred. No. 55;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||| |||  
Db 48 KLVVIAED 55

#### RESULT 84

RL7A\_THEVO  
ID RL7A\_THEVO STANDARD; PRT; 121 AA.  
AC Q97BK8;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 50S ribosomal protein L7Ae.  
 GN RPL7AE OR TV0447 OR TVG0437556.  
 OS Thermoplasma volcanium.  
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
 OC Thermoplasmataceae; Thermoplasma.  
 OX NCBI\_TaxID=50339;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GSS1 / DSM 4299 / JCM 9571;  
 RX MEDLINE=20570466; PubMed=11121031;  
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
 RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
 RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
 RT "Archaeal adaptation to higher temperatures revealed by genomic  
 RT sequence of Thermoplasma volcanium."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
 CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP000992; BAB59589.1; -.  
 DR HAMAP; MF\_00326; -; 1.  
 DR InterPro; IPR004038; Ribosomal\_L7A.  
 DR InterPro; IPR004037; Ribosomal\_L7Ae.  
 DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
 DR PRINTS; PR00881; L7ARS6FAMILY.  
 DR PROSITE; PS01082; RIBOSOMAL\_L7AE; FALSE\_NEG.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 121 AA; 13389 MW; 19FE3A349D3F9D36 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 121;  
 Best Local Similarity 75.0%; Pred. No. 55;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||| |||  
 Db 48 KLVVIAED 55

#### RESULT 85

RL7A\_AERPE

ID RL7A\_AERPE STANDARD; PRT; 127 AA.  
 AC Q9YAX7;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L7Ae.  
 GN RPL7AE OR APE1818.  
 OS Aeropyrum pernix.  
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
 OC Desulfurococcaceae; Aeropyrum.

OX NCBI\_TaxID=56636;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,  
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 RT crenarchaeon, *Aeropyrum pernix* K1.";  
 RL DNA Res. 6:83-101(1999).  
 CC -!- SIMILARITY: Belongs to the L7AE family of ribosomal proteins.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP000062; BAA80821.1; -.  
 DR PIR; H72566; H72566.  
 DR HSSP; P55769; 1E7K.  
 DR HAMAP; MF\_00326; -; 1.  
 DR InterPro; IPR004038; Ribosomal\_L7A.  
 DR InterPro; IPR004037; Ribosomal\_L7Ae.  
 DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
 DR PRINTS; PR00881; L7ARS6FAMILY.  
 DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 127 AA; 13952 MW; 21BF3E1A072BB8E6 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 127;  
 Best Local Similarity 75.0%; Pred. No. 58;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||| |||  
 Db 49 KLVVIAED 56

RESULT 86  
 YUFK\_BACSU  
 ID YUFK\_BACSU STANDARD; PRT; 204 AA.  
 AC O05249;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical protein yufK.  
 GN YUFK OR BSU31510.  
 OS *Bacillus subtilis*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.  
 OX NCBI\_TaxID=1423;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=97419515; PubMed=9274030;  
RA Oudega B., Koningstein G., Rodrigues L., de Sales Ramon M.,  
RA Hilbert H., Duesterhoeft A., Pohl T.M., Weitzenegger T.;  
RT "Analysis of the Bacillus subtilis genome: cloning and nucleotide  
RT sequence of a 62 kb region between 275 degrees (rrnB) and 284 degrees  
RT (pai).";  
RL Microbiology 143:2769-2774(1997).

RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis.";  
RL Nature 390:249-256(1997).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; Z93937; CAB07945.1; -.  
DR EMBL; Z99120; CAB15140.1; -.

DR PIR; H70008; H70008.  
 DR SubtiList; BG12346; yufK.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 28 50 Potential.  
 FT TRANSMEM 70 92 Potential.  
 FT TRANSMEM 170 192 Potential.  
 SQ SEQUENCE 204 AA; 23235 MW; 1EF138BDE6B72CF0 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 204;  
 Best Local Similarity 71.4%; Pred. No. 92;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LVFFAED 8  
 |:| | | |  
 Db 102 LLFFAND 108

# RESULT 87

## TATC\_ECOLI

ID TATC\_ECOLI STANDARD; PRT; 258 AA.  
 AC P27857; P27858; P76765; P76766;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Sec-independent protein translocase protein tatC.  
 GN TATC OR MTTB OR B3839 OR Z5360 OR ECS4768.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562, 83334;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=K12;  
 RX MEDLINE=98315056; PubMed=9649434;  
 RA Sargent F., Bogsch E.G., Stanley N.R., Wexler M., Robinson C.,  
 RA Berks B.C., Palmer T.;  
 RT "Overlapping functions of components of a bacterial Sec-independent  
 RT protein export pathway."  
 RL EMBO J. 17:3640-3650(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33694 / HB101;  
 RX MEDLINE=98206471; PubMed=9546395;  
 RA Weiner J.H., Bilous P.T., Shaw G.M., Lubitz S.P., Frost L.,  
 RA Thomas G.H., Cole J.A., Turner R.J.;  
 RT "A novel and ubiquitous system for membrane targeting and secretion of  
 RT cofactor-containing proteins."  
 RL Cell 93:93-101(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=92358234; PubMed=1379743;  
 RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;  
 RT "Analysis of the Escherichia coli genome: DNA sequence of the region  
 RT from 84.5 to 86.5 minutes."  
 RL Science 257:771-778(1992).

RN [4]  
 RP REVISIONS.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 RN [7]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98324996; PubMed=9660752;  
 RA Bogsch E.G., Sargent F., Stanley N.R., Berks B.C., Robinson C.,  
 RA Palmer T.;  
 RT "An essential component of a novel bacterial protein export system  
 RT with homologues in plastids and mitochondria.";  
 RL J. Biol. Chem. 273:18003-18006(1998).  
 RN [8]  
 RP REVIEW.  
 RX MEDLINE=20117987; PubMed=10652088;  
 RA Berks B.C., Sargent F., Palmer T.;  
 RT "The Tat protein export pathway.";  
 RL Mol. Microbiol. 35:260-274(2000).  
 CC -!- FUNCTION: Required for correct localization of precursor proteins  
 CC bearing signal peptides with the twin arginine conserved motif  
 CC S/T-R-R-X-F-L-K. This sec-independent pathway is termed TAT for  
 CC twin-arginine translocation system. This system mainly transports  
 CC proteins with bound cofactors that require folding prior to  
 CC export.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Probable).  
 CC -!- SIMILARITY: Belongs to the tatC family.  
 CC -----



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AJ005830; CAA06726.1; -.  
DR EMBL; AF067848; AAC19242.1; -.  
DR EMBL; M87049; AAA67634.1; ALT\_FRAME.  
DR EMBL; M87049; AAA67635.1; ALT\_FRAME.  
DR EMBL; AE000459; AAC76842.1; -.  
DR EMBL; AE005615; AAG59034.1; -.  
DR EMBL; AP002567; BAB38191.1; -.  
DR PIR; F86071; F86071.  
DR PIR; H65188; H65188.  
DR PIR; H91224; H91224.  
DR PIR; S30728; S30728.  
DR EcoGene; EG11479; tatC.  
DR InterPro; IPR008277; Tat\_C.  
DR InterPro; IPR002033; Translocase.  
DR Pfam; PF00902; TatC; 1.  
DR TIGRFAMs; TIGR00945; tatC; 1.  
DR PROSITE; PS01218; TATC; 1.  
KW Transport; Protein transport; Translocation; Transmembrane;  
KW Inner membrane; Complete proteome.  
FT TRANSMEM 24 44 POTENTIAL.  
FT TRANSMEM 76 96 POTENTIAL.  
FT TRANSMEM 116 136 POTENTIAL.  
FT TRANSMEM 157 177 POTENTIAL.  
FT TRANSMEM 194 210 POTENTIAL.  
FT TRANSMEM 212 232 POTENTIAL.  
SQ SEQUENCE 258 AA; 28876 MW; D2383F85AF62A81C CRC64;

Query Match 67.5%; Score 27; DB 1; Length 258;  
Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
||:|||  
Db 34 LVYFAND 40

RESULT 88

YDJO\_ECOLI

ID YDJO\_ECOLI STANDARD; PRT; 271 AA.  
AC P76210;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein ydjO.  
GN YDJO OR B1730.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12."  
 RL Science 277:1453-1474(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000268; AAC74800.1; -.  
 DR PIR; B64932; B64932.  
 DR EcoGene; EG13991; ydjO.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 271 AA; 30732 MW; BF1CC2656129AD4A CRC64;

Query Match 67.5%; Score 27; DB 1; Length 271;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
 :|||:|  
 Db 236 IFFADD 241

# RESULT 89

UPK\_CHLTE  
 ID UPK\_CHLTE STANDARD; PRT; 282 AA.  
 AC Q8KFJ7;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative undecaprenol kinase (EC 2.7.1.66) (Bacitracin resistance  
 DE protein).  
 GN UPK OR BACA OR CT0329.  
 OS Chlorobium tepidum.  
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
 OC Chlorobium.  
 OX NCBI\_TaxID=1097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TLS / ATCC 49652 / DSM 12025;  
 RX MEDLINE=22103685; PubMed=12093901;  
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,  
 RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,  
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,  
 RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,  
 RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,

RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,  
 RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;  
 RT "The complete genome sequence of *Chlorobium tepidum* TLS, a  
 RT photosynthetic, anaerobic, green-sulfur bacterium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).  
 CC -!- FUNCTION: Probably phosphorylates undecaprenol to undecaprenyl  
 CC phosphate. Confers resistance to bacitracin (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + undecaprenol = ADP + undecaprenyl  
 CC phosphate.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -!- MISCELLANEOUS: Bacitracin is thought to be involved in inhibition  
 CC of peptidoglycan synthesis by sequestering undecaprenyl  
 CC diphosphate reducing the pool of lipid carrier available.  
 CC -!- SIMILARITY: Belongs to the upk family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE012811; AAM71575.1; -.  
 DR TIGR; CT0329; -.  
 DR HAMAP; MF\_01006; -; 1.  
 DR InterPro; IPR003824; BacA.  
 DR Pfam; PF02673; BacA; 1.  
 DR TIGRFAMs; TIGR00753; undec\_kin\_bacA; 1.  
 KW Transferase; Kinase; Antibiotic resistance; Transmembrane;  
 KW Complete proteome.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 88 105 POTENTIAL.  
 FT TRANSMEM 115 132 POTENTIAL.  
 FT TRANSMEM 196 215 POTENTIAL.  
 FT TRANSMEM 230 252 POTENTIAL.  
 FT TRANSMEM 259 281 POTENTIAL.  
 SQ SEQUENCE 282 AA; 30196 MW; 8231A028E2C17C19 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 282;  
 Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
 |::||:|  
 Db 56 LIYFAKD 62

RESULT 90  
 UPK\_CORGL  
 ID UPK\_CORGL STANDARD; PRT; 293 AA.  
 AC Q8NQC3;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative undecaprenol kinase (EC 2.7.1.66) (Bacitracin resistance  
 DE protein).

GN UPK OR BACA OR CGL1515.  
 OS *Corynebacterium glutamicum* (*Brevibacterium flavum*).  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=1718;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
 RA Nakagawa S.;  
 RT "Complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Probably phosphorylates undecaprenol to undecaprenyl  
 CC phosphate. Confers resistance to bacitracin (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + undecaprenol = ADP + undecaprenyl  
 CC phosphate.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -!- MISCELLANEOUS: Bacitracin is thought to be involved in inhibition  
 CC of peptidoglycan synthesis by sequestering undecaprenyl  
 CC diphosphate reducing the pool of lipid carrier available.  
 CC -!- SIMILARITY: Belongs to the upk family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP005278; BAB98908.1; -.  
 DR HAMAP; MF\_01006; -; 1.  
 DR InterPro; IPR003824; BacA.  
 DR Pfam; PF02673; BacA; 1.  
 DR TIGRFAMs; TIGR00753; undec\_kin\_bacA; 1.  
 KW Transferase; Kinase; Antibiotic resistance; Transmembrane;  
 KW Complete proteome.  
 FT TRANSMEM 72 94 POTENTIAL.  
 FT TRANSMEM 101 123 POTENTIAL.  
 FT TRANSMEM 133 151 POTENTIAL.  
 FT TRANSMEM 207 229 POTENTIAL.  
 FT TRANSMEM 239 261 POTENTIAL.  
 FT TRANSMEM 268 290 POTENTIAL.  
 SQ SEQUENCE 293 AA; 31578 MW; 8FCD9F77D16D136E CRC64;

Query Match 67.5%; Score 27; DB 1; Length 293;  
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
 |||||::  
 Db 75 LVFFAKE 81

RESULT 91  
 XT31\_ARATH  
 ID XT31\_ARATH STANDARD; PRT; 293 AA.  
 AC P93046; Q9LXH6;

DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Probable xyloglucan endotransglucosylase/hydrolase protein 31  
 DE precursor (EC 2.4.1.207) (At-XTH31) (XTH-31) (AtXTR8).  
 GN XTH31 OR ATXG OR XTR8 OR AT3G44990 OR F14D17.60.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND  
 RP INDUCTION.  
 RC TISSUE=Seed;  
 RA Aubert D., Herzog M.;  
 RT "A new cDNA encoding a xyloglucan endo-transglycosylase-related  
 RT polypeptide (AtXTR8) preferentially expressed in seedling, root and  
 RT stem of Arabidopsis thaliana.";  
 RL Plant Sci. 121:187-196(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016720; PubMed=11130713;  
 RA Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,  
 RA Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,  
 RA Delsen Y., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,  
 RA De Simone V., Choise N., Artiguenave F., Robert C., Brottier P.,  
 RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,  
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,  
 RA Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,  
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,  
 RA Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,  
 RA Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,  
 RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,  
 RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,  
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,  
 RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,  
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,  
 RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
 RA Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,  
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,  
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
 RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,  
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:820-822(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172;

RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,  
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome.";  
 RL Science 302:842-846(2003).  
 RN [4]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=21530286; PubMed=11673616;  
 RA Yokoyama R., Nishitani K.;  
 RT "A comprehensive expression analysis of all members of a gene family  
 RT encoding cell-wall enzymes allowed us to predict cis-regulatory  
 RT regions involved in cell-wall construction in specific organs of  
 RT Arabidopsis.";  
 RL Plant Cell Physiol. 42:1025-1033(2001).  
 RN [5]  
 RP NOMENCLATURE.  
 RX MEDLINE=22402747; PubMed=12514239;  
 RA Rose J.K.C., Braam J., Fry S.C., Nishitani K.;  
 RT "The XTH family of enzymes involved in xyloglucan  
 RT endotransglucosylation and endohydrolysis: current perspectives and a  
 RT new unifying nomenclature.";  
 RL Plant Cell Physiol. 43:1421-1435(2002).  
 CC -!- FUNCTION: Catalyzes xyloglucan endohydrolysis (XEH) and/or  
 CC endotransglycosylation (XET). Cleaves and religates xyloglucan  
 CC polymers, an essential constituent of the primary cell wall, and  
 CC thereby participates in cell wall construction of growing tissues  
 CC (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Breaks a beta-(1->4) bond in the backbone of a  
 CC xyloglucan and transfers the xyloglucanyl segment on to O-4 of the  
 CC non-reducing terminal glucose residue of an acceptor, which can be  
 CC a xyloglucan or an oligosaccharide of xyloglucan.  
 CC -!- SUBCELLULAR LOCATION: Apoplast (Probable).  
 CC -!- TISSUE SPECIFICITY: Predominantly expressed in root. Weakly  
 CC expressed in inflorescence stems.  
 CC -!- DEVELOPMENTAL STAGE: Expressed in germinating seeds from 24 hours  
 CC after imbibition, and reaches a maximum level at 72 hours. After  
 CC 96 hours, it then decreases.  
 CC -!- INDUCTION: By gibberellins (Probable). Not induced by auxin.  
 CC -!- PTM: Contains at least one intrachain disulfide bond essential for  
 CC its enzymatic activity (By similarity).  
 CC -!- MISCELLANEOUS: In contrast to group 1 and group 2  
 CC endotransglucosylase/hydrolase proteins, it may not contain the  
 CC ligase activity, and may catalyze endohydrolysis xyloglucan  
 CC polymers only.  
 CC -!- SIMILARITY: Belongs to family 16 of glycosyl hydrolases. XTH  
 CC group 3 subfamily.  
 CC -!- DATABASE: NAME=XTH-World;

CC WWW="http://www.plantbio.cornell.edu/XTH".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X92975; CAA63553.1; -.  
 DR EMBL; AL353992; CAB89314.1; -.  
 DR EMBL; AY056163; AAL07012.1; -.  
 DR EMBL; AY136454; AAM97119.1; -.  
 DR EMBL; BT006326; AAP13434.1; -.  
 DR PIR; T48975; T48975.  
 DR HSSP; P23904; 1AJK.  
 DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
 DR InterPro; IPR000757; Glyco\_hydro\_16.  
 DR Pfam; PF00722; Glyco\_hydro\_16; 1.  
 DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; FALSE\_NEG.  
 KW Hydrolase; Transferase; Glycosidase; Cell wall; Apoplast; Signal;  
 KW Multigene family.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 293 PROBABLE XYLOGLUCAN  
 FT ENDOTRANSGLUCOSYLASE/HYDROLASE PROTEIN  
 FT 31.  
 FT ACT\_SITE 114 114 NUCLEOPHILE (BY SIMILARITY).  
 FT ACT\_SITE 118 118 PROTON DONOR (BY SIMILARITY).  
 FT CONFLICT 11 11 L -> V (IN REF. 1).  
 FT CONFLICT 260 260 S -> T (IN REF. 1).  
 SQ SEQUENCE 293 AA; 33540 MW; D999B4C62C5ABCA8 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 293;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 ::||| :|  
 Db 172 QIVFFVDD 179

# RESULT 92

## XTH5\_ARATH

ID XTH5\_ARATH STANDARD; PRT; 293 AA.  
 AC Q9XIW1;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Probable xyloglucan endotransglucosylase/hydrolase protein 5 precursor  
 DE (EC 2.4.1.207) (At-XTH5) (XTH-5).  
 GN XTH5 OR EXGT-A4 OR XTR12 OR AT5G13870 OR MAC12.33.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;

RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20027412; PubMed=10557219;  
 RA Akamatsu T., Hanzawa Y., Ohtake Y., Takahashi T., Nishitani K.,  
 RA Komeda Y.;  
 RT "Expression of endoxyloglucan transferase genes in acaulis mutants of  
 RT Arabidopsis.";  
 RL Plant Physiol. 121:715-721(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Okamoto S., Hayashida N., Shinozaki K., Nishitani K.;  
 RT "Endoxyloglucan transferase homologue from Arabidopsis thaliana.";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=97471969; PubMed=9330910;  
 RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,  
 RA Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence  
 RT features of the 1.6 Mb regions covered by twenty physically assigned  
 RT Pl clones.";  
 RL DNA Res. 4:215-230(1997).  
 RN [4]  
 RP TISSUE SPECIFICITY, AND INDUCTION.  
 RX MEDLINE=21530286; PubMed=11673616;  
 RA Yokoyama R., Nishitani K.;  
 RT "A comprehensive expression analysis of all members of a gene family  
 RT encoding cell-wall enzymes allowed us to predict cis-regulatory  
 RT regions involved in cell-wall construction in specific organs of  
 RT Arabidopsis.";  
 RL Plant Cell Physiol. 42:1025-1033(2001).  
 RN [5]  
 RP NOMENCLATURE.  
 RX MEDLINE=22402747; PubMed=12514239;  
 RA Rose J.K.C., Braam J., Fry S.C., Nishitani K.;  
 RT "The XTH family of enzymes involved in xyloglucan  
 RT endotransglucosylation and endohydrolysis: current perspectives and a  
 RT new unifying nomenclature.";  
 RL Plant Cell Physiol. 43:1421-1435(2002).  
 CC -!- FUNCTION: Catalyzes xyloglucan endohydrolysis (XEH) and/or  
 CC endotransglycosylation (XET). Cleaves and religates xyloglucan  
 CC polymers, an essential constituent of the primary cell wall, and  
 CC thereby participates in cell wall construction of growing tissues  
 CC (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Breaks a beta-(1->4) bond in the backbone of a  
 CC xyloglucan and transfers the xyloglucanyl segment on to O-4 of the  
 CC non-reducing terminal glucose residue of an acceptor, which can be  
 CC a xyloglucan or an oligosaccharide of xyloglucan.  
 CC -!- SUBCELLULAR LOCATION: Apoplast (Probable).  
 CC -!- TISSUE SPECIFICITY: Root specific.  
 CC -!- INDUCTION: By brassinolide. Strongly downregulated by abscisic  
 CC acid.  
 CC -!- PTM: Contains at least one intrachain disulfide bond essential for  
 CC its enzymatic activity (By similarity).  
 CC -!- SIMILARITY: Belongs to family 16 of glycosyl hydrolases. XTH



CC       goup 1 subfamily.  
 CC    -!- DATABASE: NAME=XTH-World;  
 CC       WWW="http://www.plantbio.cornell.edu/XTH".  
 CC    -----  
 CC    This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC    the European Bioinformatics Institute. There are no restrictions on its  
 CC    use by non-profit institutions as long as its content is in no way  
 CC    modified and this statement is not removed. Usage by and for commercial  
 CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC    -----  
 DR    EMBL; AF163822; AAD45126.1; -.  
 DR    EMBL; AB026486; BAA81669.1; -.  
 DR    EMBL; AB005230; BAB11115.1; -.  
 DR    HSSP; P23904; 1AJK.  
 DR    InterPro; IPR008264; Beta\_glucanase.  
 DR    InterPro; IPR008985; ConA\_like\_lec\_gl.  
 DR    InterPro; IPR008263; Glyco\_hyd16\_AS.  
 DR    InterPro; IPR000757; Glyco\_hydro\_16.  
 DR    Pfam; PF00722; Glyco\_hydro\_16; 1.  
 DR    PRINTS; PR00737; GLHYDRLASE16.  
 DR    PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
 KW    Hydrolase; Transferase; Glycosidase; Cell wall; Apoplast; Signal;  
 KW    Multigene family.  
 FT    SIGNAL           1       21       POTENTIAL.  
 FT    CHAIN           22       293      PROBABLE XYLOGLUCAN  
 FT                                    ENDOTRANSGLUCOSYLASE/HYDROLASE PROTEIN 5.  
 FT    ACT\_SITE       106       106      NUCLEOPHILE (BY SIMILARITY).  
 FT    ACT\_SITE       110       110      PROTON DONOR (BY SIMILARITY).  
 FT    CARBOHYD       114       114      N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ    SEQUENCE       293 AA;   33949 MW;   1025767DDA2BA9F4 CRC64;

Query Match                   67.5%;   Score 27;   DB 1;   Length 293;  
 Best Local Similarity       50.0%;   Pred. No. 1.3e+02;  
 Matches       4;   Conservative       3;   Mismatches       1;   Indels       0;   Gaps       0;

Qy           1 KLVFFAED 8  
              ::||| :|  
 Db           159 QIVFFVDD 166

RESULT 93  
 Y347\_HELPJ  
 ID   Y347\_HELPJ       STANDARD;       PRT;   293 AA.  
 AC   Q9ZMA1;  
 DT   16-OCT-2001 (Rel. 40, Created)  
 DT   16-OCT-2001 (Rel. 40, Last sequence update)  
 DT   10-OCT-2003 (Rel. 42, Last annotation update)  
 DE   Hypothetical pseudouridine synthase JHP0321 (EC 4.2.1.70)  
 DE   (Pseudouridylate synthase) (Uracil hydrolyase).  
 GN   JHP0321.  
 OS   Helicobacter pylori J99 (Campylobacter pylori J99).  
 OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC   Helicobacteraceae; Helicobacter.  
 OX   NCBI\_TaxID=85963;  
 RN   [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=99120557; PubMed=9923682;  
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
RA Trust T.J.;  
RT "Genomic sequence comparison of two unrelated isolates of the human  
RT gastric pathogen *Helicobacter pylori*.";  
RL Nature 397:176-180(1999).  
CC -!- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine  
CC 5'-phosphate + H(2)O.  
CC -!- SIMILARITY: Belongs to the pseudouridine synthase rluA family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE001468; AAD05911.1; -.  
DR PIR; A71946; A71946.  
DR InterPro; IPR006145; PseudoU\_synth.  
DR InterPro; IPR006224; Rlu\_synth.  
DR Pfam; PF00849; PseudoU\_synth\_2; 1.  
DR ProDom; PD001819; PSI\_RLU; 1.  
DR PROSITE; PS01129; PSI\_RLU; FALSE\_NEG.  
KW Hypothetical protein; Lyase; Complete proteome.  
FT ACT\_SITE 119 119 BY SIMILARITY.  
SQ SEQUENCE 293 AA; 33668 MW; D7EF3EC54A8E6168 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 293;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||| | |:  
Db 70 KLVFEAKD 77

RESULT 94  
GBB\_SCHPO  
ID GBB\_SCHPO STANDARD; PRT; 305 AA.  
AC Q10282;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Guanine nucleotide-binding protein beta subunit.  
GN GIT5 OR GPB1 OR SPBC32H8.07 OR PI017.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=96397506; PubMed=8804400;  
 RA Kim D.U., Park S.K., Chung K.S., Choi M.U., Yoo H.S.;  
 RT "The G protein beta subunit Gpb1 of Schizosaccharomyces pombe is a  
 RT negative regulator of sexual development.";  
 RL Mol. Gen. Genet. 252:20-32(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20211472; PubMed=10747045;  
 RA Landry S., Pettit M.T., Apolinario E., Hoffman C.S.;  
 RT "The fission yeast git5 gene encodes a gbeta subunit required for  
 RT glucose-triggered adenylate cyclase activation.";  
 RL Genetics 154:1463-1471(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=20089027; PubMed=10620777;  
 RA Machida M., Yamazaki S., Kunihiro S., Tanaka T., Kushida N., Jinno K.,  
 RA Haikawa Y., Yamazaki J., Yamamoto S., Sekine M., Oguchi A., Nagai Y.,  
 RA Sakai M., Aoki K., Ogura K., Kudoh Y., Kikuchi H., Zhang M.Q.,  
 RA Yanagida M.;  
 RT "A 38 kb segment containing the cdc2 gene from the left arm of fission  
 RT yeast chromosome II: sequence analysis and characterization of the  
 RT genomic DNA and cDNAs encoded on the segment.";  
 RL Yeast 16:71-80(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 CC INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE

CC SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE  
 CC GTPASE ACTIVITY, FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-  
 CC EFFECTOR INTERACTION. REQUIRED FOR ADENYLATE CYCLASE ACTIVATION.  
 CC -!- SUBUNIT: G proteins are composed of 3 units, alpha, beta and  
 CC gamma.  
 CC -!- SIMILARITY: Contains 7 WD repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L28061; AAC37501.1; ALT\_INIT.  
 DR EMBL; AF092102; AAD09020.1; -.  
 DR EMBL; AB004535; BAA21396.1; -.  
 DR EMBL; AL590971; CAC37497.1; -.  
 DR PIR; T50474; T50474.  
 DR HSSP; P04901; 1A0R.  
 DR GeneDB\_SPombe; SPBC32H8.07; -.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; ISS.  
 DR GO; GO:0007165; P:signal transduction; ISS.  
 DR InterPro; IPR001632; Gprotein\_B.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00319; GPROTEINB.  
 DR PRINTS; PR00320; GPROTEINBRPT.  
 DR ProDom; PD000018; WD40; 1.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 2.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 5.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Transducer; Repeat; WD repeat; Sporulation.  
 FT REPEAT 19 49 WD 1.  
 FT REPEAT 61 91 WD 2.  
 FT REPEAT 104 133 WD 3.  
 FT REPEAT 145 176 WD 4.  
 FT REPEAT 188 218 WD 5.  
 FT REPEAT 231 260 WD 6.  
 FT REPEAT 272 302 WD 7.  
 SQ SEQUENCE 305 AA; 32828 MW; 7186EB505C68980E CRC64;

Query Match 67.5%; Score 27; DB 1; Length 305;  
 Best Local Similarity 62.5%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||:| |:|  
 Db 246 KLLFIAKD 253

RESULT 95  
 AHA1\_HUMAN  
 ID AHA1\_HUMAN STANDARD; PRT; 338 AA.  
 AC O95433; Q96IL6; Q9P060;

DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Activator of 90 kDa heat shock protein ATPase homolog 1 (AHA1) (p38)  
 DE (HSPC322).  
 GN AHSAL OR C14ORF3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Michaud J., Chrast R., Rossier C., Papassavas M.P., Antonarakis S.E.,  
 RA Scott H.S.;  
 RT "Isolation of a novel gene underexpressed in Down syndrome.";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Adrenal gland;  
 RX MEDLINE=20402571; PubMed=10931946;  
 RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,  
 RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,  
 RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,  
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,  
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;  
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal  
 RT axis and full-length cDNA cloning.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung, and Ovary;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 67-338 FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=20499367; PubMed=11042152;  
 RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,

RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,  
RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;  
RT "Cloning and functional analysis of cDNAs with open reading frames for  
RT 300 previously undefined genes expressed in CD34+ hematopoietic  
RT stem/progenitor cells.";  
RL Genome Res. 10:1546-1560(2000).

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF111168; AAD09623.1; -.  
DR EMBL; AJ243310; CAB45684.1; -.  
DR EMBL; AF164791; AAF80755.1; -.  
DR EMBL; BC000321; AAH00321.1; -.  
DR EMBL; BC007398; AAH07398.1; ALT\_INIT.  
DR EMBL; AF161440; AAF29000.1; -.  
DR PIR; JC7769; JC7769.  
DR Genew; HGNC:1189; AHSA1.  
DR InterPro; IPR007821; DUF704.  
DR Pfam; PF05146; DUF704; 1.  
FT CONFLICT 67 68 EA -> CL (IN REF. 4).  
SQ SEQUENCE 338 AA; 38274 MW; E6B686DDD8D7D729 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 338;  
Best Local Similarity 71.4%; Pred. No. 1.5e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
||:|| |  
Db 76 KLIFFYE 82

#### RESULT 96

##### AHA1\_MOUSE

ID AHA1\_MOUSE STANDARD; PRT; 338 AA.  
AC Q8BK64; Q8R3E6;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Activator of 90 kDa heat shock protein ATPase homolog 1 (AHA1).  
GN AHSA1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J;  
RX MEDLINE=22354683; PubMed=12466851;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perteu G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;

RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N; TISSUE=Mammary gland;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; AK076069; BAC36160.1; -.

DR EMBL; BC023857; AAH23857.1; -.

DR EMBL; BC025552; AAH25552.1; -.

DR MGD; MGI:2387603; Ahsal.

DR InterPro; IPR007821; DUF704.

DR Pfam; PF05146; DUF704; 1.

FT CONFLICT 192 192 P -> L (IN REF. 2).

SQ SEQUENCE 338 AA; 38117 MW; C4F0F94FF44F9FB3 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 338;

Best Local Similarity 71.4%; Pred. No. 1.5e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7

||:|| |

Db 76 KLIFYE 82

#### RESULT 97

YGH5\_YEAST

ID YGH5\_YEAST STANDARD; PRT; 387 AA.

AC P53159;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Hypothetical 44.6 kDa protein in RPL7A-HSF1 intergenic region.

GN YGL075C.

OS *Saccharomyces cerevisiae* (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288c;

RX MEDLINE=97435481; PubMed=9290212;

RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;

RT "Sequence analysis of 203 kilobases from *Saccharomyces cerevisiae*

RT chromosome VII.";

RL Yeast 13:1077-1090(1997).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).

CC

DR EMBL; Z72597; CAA96780.1; -.

DR PIR; S64082; S64082.

DR GermOnline; 141123; -.

DR SGD; S0003043; MPS2.

DR GO; GO:0005635; C:nuclear membrane; IDA.

DR GO; GO:0005816; C:spindle pole body; IDA.

DR GO; GO:0005200; F:structural constituent of cytoskeleton; IMP.



KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 311 327 POTENTIAL.  
SQ SEQUENCE 387 AA; 44585 MW; CF8AFA86EDEA834C CRC64;

Query Match 67.5%; Score 27; DB 1; Length 387;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|:|:| ||  
Db 356 KIVWFFED 363

RESULT 98

YM14\_PARTE

ID YM14\_PARTE STANDARD; PRT; 387 AA.  
AC P15615;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-APR-1990 (Rel. 14, Last annotation update)  
DE Hypothetical 47.2 kDa protein (ORF14).  
OS Paramecium tetraurelia.  
OG Mitochondrion.  
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;  
OC Paramecium.  
OX NCBI\_TaxID=5888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Stock 51;  
RX MEDLINE=90174913; PubMed=2308823;  
RA Pritchard A.E., Seilhamer J.J., Mahalingam R., Sable C.L.,  
RA Venuti S.E., Cummings D.J.;  
RT "Nucleotide sequence of the mitochondrial genome of Paramecium."  
RL Nucleic Acids Res. 18:173-180(1990).

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X15917; CAA34028.1; -.  
DR PIR; S07749; S07749.  
KW Hypothetical protein; Mitochondrion.  
SQ SEQUENCE 387 AA; 47282 MW; AEEBC6315F4F1CAD CRC64;

Query Match 67.5%; Score 27; DB 1; Length 387;  
Best Local Similarity 71.4%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
||:|| |  
Db 34 KLIFFFE 40

RESULT 99

AV71\_ACAVI

ID AV71\_ACAVI STANDARD; PRT; 394 AA.  
AC Q17107;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Muscle cell intermediate filament protein AV71 (Fragment).  
GN AV71.  
OS Acanthocheilonema viteae (Filarial nematode) (Dipetalonema viteae).  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
OC Onchocercidae; Acanthocheilonema.  
OX NCBI\_TaxID=6277;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Muscle;  
RX MEDLINE=95192264; PubMed=7886042;  
RA Seeber F., Hoefle W., Kern A., Lucius R.;  
RT "Onchocerca volvulus and Acanthocheilonema viteae: cloning of cDNAs  
RT for muscle-cell intermediate filaments.";  
RL Parasitol. Res. 80:699-702(1994).  
CC -!- SIMILARITY: Belongs to the intermediate filament family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X68557; CAA48560.1; -.  
DR PIR; S26431; S26431.  
DR InterPro; IPR001664; IF.  
DR InterPro; IPR001322; IF\_tail\_C.  
DR Pfam; PF00038; filament; 1.  
DR Pfam; PF00932; IF\_tail; 1.  
DR PROSITE; PS00226; IF; 1.  
KW Intermediate filament; Coiled coil.  
FT NON\_TER 1 1  
FT DOMAIN <1 239 ROD.  
FT DOMAIN 240 394 TAIL.  
FT DOMAIN <1 73 COIL 1B.  
FT DOMAIN 74 91 LINKER 12.  
FT DOMAIN 92 239 COIL 2.  
SQ SEQUENCE 394 AA; 45862 MW; 5508ECE650DF9354 CRC64;

Query Match 67.5%; Score 27; DB 1; Length 394;  
Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
:| | | |  
Db 355 QLVFDAED 362

RESULT 100

OV71\_ONCVO

ID OV71\_ONCVO STANDARD; PRT; 432 AA.

AC P31732;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Muscle cell intermediate filament protein OV71 (Fragment).

GN OV71.

OS Onchocerca volvulus.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;

OC Onchocercidae; Onchocerca.

OX NCBI\_TaxID=6282;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Muscle;

RX MEDLINE=95192264; PubMed=7886042;

RA Seeber F., Hoefle W., Kern A., Lucius R.;

RT "Onchocerca volvulus and Acanthocheilonema viteae: cloning of cDNAs

RT for muscle-cell intermediate filaments.";

RL Parasitol. Res. 80:699-702(1994).

CC -!- SIMILARITY: Belongs to the intermediate filament family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X68558; CAA48561.1; -.

DR PIR; S26432; S26432.

DR InterPro; IPR001664; IF.

DR InterPro; IPR001322; IF\_tail\_C.

DR Pfam; PF00038; filament; 1.

DR Pfam; PF00932; IF\_tail; 1.

DR PROSITE; PS00226; IF; 1.

KW Intermediate filament; Coiled coil.

FT NON\_TER 1 1

FT DOMAIN <1 277 ROD.

FT DOMAIN 278 432 TAIL.

FT DOMAIN <1 111 COIL 1B.

FT DOMAIN 112 128 LINKER 12.

FT DOMAIN 129 277 COIL 2.

SQ SEQUENCE 432 AA; 50399 MW; 9E2824B2D3DFD55D CRC64;

Query Match 67.5%; Score 27; DB 1; Length 432;

Best Local Similarity 75.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8

:||| |||

Db 393 QLVFDAED 400

Search completed: February 28, 2004, 08:52:52

Job time : 66.5 secs

OM protein - protein search, using sw model

Run on: February 28, 2004, 08:36:39 ; Search time 60.5 Seconds  
(without alignments)  
41.721 Million cell updates/sec

Title: US-09-668-314C-73  
Perfect score: 40  
Sequence: 1 KLVFFAED 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	%	Query					
No.	Score	Match	Length	DB	ID		Description
-----							

1	40	100.0	28	4	Q9UCD1	Q9ucd1 homo sapien
2	40	100.0	30	4	Q9UCA9	Q9uca9 homo sapien
3	40	100.0	33	4	Q9UC33	Q9uc33 homo sapien
4	40	100.0	79	11	O35463	O35463 cricetulus
5	40	100.0	82	4	Q16020	Q16020 homo sapien
6	40	100.0	82	4	Q16014	Q16014 homo sapien
7	40	100.0	82	4	Q16019	Q16019 homo sapien
8	40	100.0	113	13	Q8JH58	Q8jh58 chelydra se
9	40	100.0	218	11	Q8BPV5	Q8bpv5 mus musculu
10	40	100.0	357	13	Q8UUI8	Q8uui8 brachydanio
11	40	100.0	384	11	Q8BPC7	Q8bpc7 mus musculu
12	40	100.0	472	13	Q8UUS0	Q8uus0 brachydanio
13	40	100.0	534	13	O93296	O93296 gallus gall
14	40	100.0	569	13	Q9PVL1	Q9pvl1 gallus gall
15	40	100.0	612	13	Q9I9E7	Q9i9e7 brachydanio
16	40	100.0	678	13	Q7ZZT1	Q7zzt1 brachydanio
17	40	100.0	695	13	Q9DGJ8	Q9dgj8 gallus gall
18	40	100.0	738	13	Q90W28	Q90w28 brachydanio
19	40	100.0	751	13	Q9DGJ7	Q9dgj7 gallus gall
20	37	92.5	1676	16	Q8A6R7	Q8a6r7 bacteroides
21	36	90.0	693	13	Q98SG0	Q98sg0 xenopus lae
22	36	90.0	747	13	Q91963	Q91963 xenopus. ap
23	33	82.5	162	8	Q32406	Q32406 heteranther
24	33	82.5	197	16	Q7VR77	Q7vr77 candidatus
25	33	82.5	695	13	Q98SF9	Q98sf9 xenopus lae
26	33	82.5	695	13	Q7ZXQ0	Q7zxq0 xenopus lae
27	32	80.0	182	16	Q9Z588	Q9z588 streptomyce
28	32	80.0	184	16	Q931V3	Q931v3 staphylococ
29	32	80.0	261	2	Q7X225	Q7x225 staphylococ
30	32	80.0	261	2	Q7WRM0	Q7wrm0 staphylococ
31	32	80.0	261	16	Q99V89	Q99v89 staphylococ
32	32	80.0	268	16	Q8NXD0	Q8nxd0 staphylococ
33	32	80.0	282	16	Q8CUH9	Q8cuh9 oceanobacil
34	32	80.0	472	13	Q10833	Q10833 xenopus lae
35	32	80.0	501	16	Q7UZT9	Q7uzt9 prochloroco
36	32	80.0	1105	5	Q9VX31	Q9vx31 drosophila
37	32	80.0	2613	5	Q9GYD1	Q9gyd1 leishmania
38	31	77.5	49	6	O97917	O97917 bos taurus
39	31	77.5	147	16	Q8A5K5	Q8a5k5 bacteroides
40	31	77.5	179	16	Q82JK4	Q82jk4 streptomyce
41	31	77.5	208	16	Q8ESR7	Q8esr7 oceanobacil
42	31	77.5	228	16	Q8E2V5	Q8e2v5 streptococc
43	31	77.5	228	16	Q8DX05	Q8dx05 streptococc
44	31	77.5	248	5	Q8I3W8	Q8i3w8 plasmodium
45	31	77.5	259	16	Q8NN32	Q8nn32 corynebacte
46	31	77.5	290	12	Q9IK76	Q9ik76 spodoptera
47	31	77.5	355	16	Q814E2	Q814e2 bacillus ce
48	31	77.5	370	16	Q7WD43	Q7wd43 bordetella
49	31	77.5	370	16	Q7W5K7	Q7w5k7 bordetella
50	31	77.5	370	16	Q7VW49	Q7vw49 bordetella
51	31	77.5	699	13	O57394	O57394 narke japon
52	31	77.5	884	16	Q83E69	Q83e69 coxiella bu
53	31	77.5	906	16	Q8AA21	Q8aa21 bacteroides
54	31	77.5	1339	3	Q9UUG9	Q9uug9 schizosacch
55	31	77.5	1364	3	Q8NIT0	Q8nit0 neurospora
56	30	75.0	92	10	Q8H8K8	Q8h8k8 oryza sativ
57	30	75.0	122	6	O77630	O77630 bos taurus

58	30	75.0	125	16	Q8CSR5	Q8csr5 staphylococ
59	30	75.0	204	16	Q931Y2	Q931y2 staphylococ
60	30	75.0	222	5	Q21915	Q21915 caenorhabdi
61	30	75.0	258	16	Q8ZAM4	Q8zam4 yersinia pe
62	30	75.0	354	16	Q898W5	Q898w5 clostridium
63	30	75.0	364	2	Q93CU7	Q93cu7 shigella bo
64	30	75.0	370	4	Q8NEN2	Q8nen2 homo sapien
65	30	75.0	441	16	Q8ZNH0	Q8znh0 salmonella
66	30	75.0	502	5	O62511	O62511 caenorhabdi
67	30	75.0	513	16	Q8A226	Q8a226 bacteroides
68	30	75.0	607	17	Q8Q077	Q8q077 methanosarc
69	30	75.0	681	3	O60157	O60157 schizosacch
70	30	75.0	740	3	Q03455	Q03455 saccharomyc
71	30	75.0	748	5	O77066	O77066 choristoneu
72	30	75.0	768	10	Q9LFI1	Q9lfi1 arabidopsis
73	30	75.0	810	3	Q9P852	Q9p852 candida alb
74	29	72.5	99	16	Q97S58	Q97s58 streptococc
75	29	72.5	99	16	Q8DQV3	Q8dqv3 streptococc
76	29	72.5	100	16	Q8Y8G0	Q8y8g0 listeria mo
77	29	72.5	101	16	Q88VK6	Q88vk6 lactobacill
78	29	72.5	119	6	Q9GMZ2	Q9gmz2 sus scrofa
79	29	72.5	137	16	Q7UPR1	Q7upr1 rhodopirell
80	29	72.5	152	10	Q9STZ9	Q9stz9 arabidopsis
81	29	72.5	160	8	Q32197	Q32197 eichhornia
82	29	72.5	162	8	Q32892	Q32892 pontederia
83	29	72.5	162	8	Q32202	Q32202 eichhornia
84	29	72.5	162	8	Q32176	Q32176 eichhornia
85	29	72.5	162	8	Q32430	Q32430 heteranther
86	29	72.5	162	8	Q32201	Q32201 eichhornia
87	29	72.5	162	8	Q32441	Q32441 heteranther
88	29	72.5	162	8	Q32802	Q32802 pontederia
89	29	72.5	162	8	Q32424	Q32424 heteranther
90	29	72.5	162	8	Q32556	Q32556 monochoria
91	29	72.5	162	8	Q32588	Q32588 monochoria
92	29	72.5	162	8	Q32390	Q32390 heteranther
93	29	72.5	162	8	Q32164	Q32164 eichhornia
94	29	72.5	162	8	Q32579	Q32579 monochoria
95	29	72.5	162	8	Q32190	Q32190 eichhornia
96	29	72.5	162	8	Q32915	Q32915 pontederia
97	29	72.5	162	8	Q32400	Q32400 hydrothrix
98	29	72.5	162	8	Q32648	Q32648 monochoria
99	29	72.5	162	8	Q32419	Q32419 heteranther
100	29	72.5	162	8	Q32179	Q32179 eichhornia
101	29	72.5	162	8	Q32210	Q32210 eichhornia
102	29	72.5	162	17	Q8PWX2	Q8pwx2 methanosarc
103	29	72.5	164	17	Q8PWX0	Q8pwx0 methanosarc
104	29	72.5	165	8	Q37011	Q37011 pontederia
105	29	72.5	168	5	Q8N0N2	Q8n0n2 littorina l
106	29	72.5	202	2	Q93P97	Q93p97 microscilla
107	29	72.5	210	16	Q8PEP5	Q8pep5 xanthomonas
108	29	72.5	213	4	Q9UJG1	Q9ujg1 homo sapien
109	29	72.5	213	11	Q8VEL0	Q8vel0 mus musculu
110	29	72.5	214	11	Q9D8Y9	Q9d8y9 mus musculu
111	29	72.5	215	10	Q94FL5	Q94fl5 arabidopsis
112	29	72.5	215	13	Q7SYQ6	Q7syq6 xenopus lae
113	29	72.5	229	10	Q8VY56	Q8vy56 arabidopsis
114	29	72.5	231	10	Q9SV79	Q9sv79 arabidopsis

115	29	72.5	235	16	Q8P3B2	Q8p3b2 xanthomonas
116	29	72.5	235	17	Q973I2	Q973i2 sulfolobus
117	29	72.5	244	16	Q897H9	Q897h9 clostridium
118	29	72.5	247	10	Q9FX75	Q9fx75 arabidopsis
119	29	72.5	273	10	Q9T0H9	Q9t0h9 arabidopsis
120	29	72.5	295	16	Q8E547	Q8e547 streptococc
121	29	72.5	295	16	Q8DZI3	Q8dzi3 streptococc
122	29	72.5	300	16	Q929N0	Q929n0 listeria in
123	29	72.5	300	16	Q8Y5C6	Q8y5c6 listeria mo
124	29	72.5	314	16	Q9CFM1	Q9cfml1 lactococcus
125	29	72.5	333	5	Q9V5C2	Q9v5c2 drosophila
126	29	72.5	333	16	Q81CC4	Q81cc4 bacillus ce
127	29	72.5	334	16	Q8R8F2	Q8r8f2 thermoanaer
128	29	72.5	335	16	Q82G15	Q82g15 streptomyce
129	29	72.5	340	5	Q8SXE5	Q8sxe5 drosophila
130	29	72.5	340	16	Q9KYD8	Q9kyd8 streptomyce
131	29	72.5	347	1	Q9P9F9	Q9p9f9 methanosarc
132	29	72.5	350	17	Q8PU59	Q8pu59 methanosarc
133	29	72.5	352	10	Q9XGY6	Q9xgy6 simmondsia
134	29	72.5	352	16	Q8YUK9	Q8yuk9 anabaena sp
135	29	72.5	354	16	Q97I04	Q97i04 clostridium
136	29	72.5	356	2	Q7WSF2	Q7wsf2 bacillus th
137	29	72.5	357	17	Q8U460	Q8u460 pyrococcus
138	29	72.5	361	8	O20025	O20025 crithmum ma
139	29	72.5	361	8	O20011	O20011 anthriscus
140	29	72.5	364	8	O20068	O20068 neogoezia m
141	29	72.5	364	8	Q8SN50	Q8sn50 lilaeopsis
142	29	72.5	364	8	Q8SN48	Q8sn48 lilaeopsis
143	29	72.5	364	8	Q8SN49	Q8sn49 lilaeopsis
144	29	72.5	366	16	Q8PPL1	Q8ppl1 xanthomonas
145	29	72.5	366	16	Q8DIZ5	Q8diz5 synechococc
146	29	72.5	369	17	Q9UZM7	Q9uzm7 pyrococcus
147	29	72.5	379	16	Q83NF3	Q83nf3 tropheryma
148	29	72.5	381	17	Q8TJA7	Q8tja7 methanosarc
149	29	72.5	389	2	Q8KHB6	Q8khb6 clostridium
150	29	72.5	390	16	Q83N16	Q83n16 tropheryma
151	29	72.5	404	16	Q8P597	Q8p597 xanthomonas
152	29	72.5	428	10	Q9M1Q8	Q9mlq8 arabidopsis
153	29	72.5	461	12	Q89329	Q89329 zucchini ye
154	29	72.5	468	10	Q94AE5	Q94ae5 arabidopsis
155	29	72.5	470	12	Q7T910	Q7t910 zucchini ye
156	29	72.5	481	4	Q86Z08	Q86z08 homo sapien
157	29	72.5	481	6	Q7YR32	Q7yr32 pan troglod
158	29	72.5	490	12	Q7T912	Q7t912 zucchini ye
159	29	72.5	490	12	Q7T911	Q7t911 zucchini ye
160	29	72.5	508	16	Q9RJ02	Q9rj02 streptomyce
161	29	72.5	558	5	Q9GP28	Q9gp28 theileria p
162	29	72.5	575	11	Q9QUU4	Q9quu4 rattus sp.
163	29	72.5	590	10	Q9M3B7	Q9m3b7 arabidopsis
164	29	72.5	607	5	Q21921	Q21921 caenorhabdi
165	29	72.5	611	6	Q8WNQ3	Q8wnq3 sus scrofa
166	29	72.5	619	16	Q8D4E8	Q8d4e8 vibrio vuln
167	29	72.5	621	16	Q9KR25	Q9kr25 vibrio chol
168	29	72.5	624	5	Q8I579	Q8i579 plasmodium
169	29	72.5	626	5	Q95ZH6	Q95zh6 cupiennius
170	29	72.5	636	16	Q83BH9	Q83bh9 coxiella bu
171	29	72.5	683	3	Q9P778	Q9p778 schizosacch

172	29	72.5	710	10	Q9SX53	Q9sx53 arabidopsis
173	29	72.5	733	4	Q8TEE2	Q8tee2 homo sapien
174	29	72.5	738	5	O97037	O97037 hydra magni
175	29	72.5	767	2	Q8KTW1	Q8ktw1 clostridium
176	29	72.5	771	11	Q80X33	Q80x33 mus musculu
177	29	72.5	816	13	Q98SV4	Q98sv4 ictalurus p
178	29	72.5	857	11	Q8K0D1	Q8k0d1 mus musculu
179	29	72.5	876	13	Q91593	Q91593 xenopus lae
180	29	72.5	884	17	Q8PW31	Q8pw31 methanosarc
181	29	72.5	898	13	Q91592	Q91592 xenopus lae
182	29	72.5	911	4	Q7Z7D0	Q7z7d0 homo sapien
183	29	72.5	955	12	Q80RZ2	Q80rz2 zantedeschi
184	29	72.5	987	8	O03684	O03684 gelasinospo
185	29	72.5	1016	12	Q05912	Q05912 zucchini ye
186	29	72.5	1028	11	P97528	P97528 rattus norv
187	29	72.5	1056	4	Q8NF42	Q8nf42 homo sapien
188	29	72.5	1077	13	Q8UWA6	Q8uwa6 tribolodon
189	29	72.5	1088	13	Q8AYH7	Q8ayh7 oncorhynchu
190	29	72.5	1097	4	Q9H6Y0	Q9h6y0 homo sapien
191	29	72.5	1165	5	Q962L9	Q962l9 plasmodium
192	29	72.5	1169	5	Q8T9D3	Q8t9d3 drosophila
193	29	72.5	1169	5	Q9VSJ6	Q9vsj6 drosophila
194	29	72.5	1340	10	Q9LTT9	Q9l9tt9 arabidopsis
195	29	72.5	1344	10	Q8L7N9	Q8l7n9 arabidopsis
196	29	72.5	1344	10	Q9LTT8	Q9l8tt8 arabidopsis
197	29	72.5	1431	5	Q8ILH5	Q8ilh5 plasmodium
198	29	72.5	1762	5	Q9BL24	Q9bl24 caenorhabdi
199	29	72.5	1866	4	Q86WI3	Q86wi3 homo sapien
200	29	72.5	1916	12	Q8QKU4	Q8qku4 zucchini ye
201	29	72.5	2079	16	Q7UF08	Q7uf08 rhodopirell
202	29	72.5	3080	12	Q7T914	Q7t914 zucchini ye
203	29	72.5	3080	12	Q7T908	Q7t908 zucchini ye
204	29	72.5	3484	5	P91257	P91257 caenorhabdi
205	29	72.5	4488	11	Q9QZH1	Q9qzh1 mus musculu
206	29	72.5	18519	5	Q8ISF6	Q8isf6 caenorhabdi
207	29	72.5	18534	5	Q8ISF7	Q8isf7 caenorhabdi
208	28	70.0	51	9	Q7Y3H8	Q7y3h8 enterobacte
209	28	70.0	59	2	O68682	O68682 bacillus ce
210	28	70.0	59	16	Q81GP5	Q81gp5 bacillus ce
211	28	70.0	98	16	Q8DJ04	Q8dj04 synechococc
212	28	70.0	117	16	Q81GV4	Q81gv4 bacillus ce
213	28	70.0	142	16	Q81U10	Q81u10 bacillus an
214	28	70.0	142	16	Q7UG51	Q7ug51 rhodopirell
215	28	70.0	169	8	Q85WI3	Q85wi3 uncultured
216	28	70.0	173	8	Q9GGR5	Q9ggr5 prasinophyc
217	28	70.0	176	3	Q08689	Q08689 saccharomyc
218	28	70.0	178	16	Q9PLW6	Q9plw6 campylobact
219	28	70.0	179	9	Q8W765	Q8w765 bacterioph
220	28	70.0	183	12	Q98243	Q98243 molluscum c
221	28	70.0	184	2	Q8KV57	Q8kv57 uncultured
222	28	70.0	184	2	Q8KVE8	Q8kve8 uncultured
223	28	70.0	184	2	Q8KV36	Q8kv36 uncultured
224	28	70.0	184	2	Q8KVF5	Q8kvf5 uncultured
225	28	70.0	184	2	Q8KUY0	Q8kuy0 uncultured
226	28	70.0	184	2	Q8KVD5	Q8kvd5 uncultured
227	28	70.0	184	2	Q8KUZ3	Q8kuz3 uncultured
228	28	70.0	184	2	Q8KVN9	Q8kvn9 uncultured



229	28	70.0	184	2	Q8KVK7	Q8kvk7 uncultured
230	28	70.0	184	2	Q8KVB2	Q8kvb2 uncultured
231	28	70.0	184	2	Q8KV30	Q8kv30 uncultured
232	28	70.0	184	2	Q8KV37	Q8kv37 uncultured
233	28	70.0	184	2	Q8KUZ6	Q8kuz6 uncultured
234	28	70.0	184	2	Q8KVD9	Q8kvd9 uncultured
235	28	70.0	184	2	Q8KI37	Q8ki37 uncultured
236	28	70.0	184	2	Q8KV39	Q8kv39 uncultured
237	28	70.0	184	2	Q8KHI1	Q8khi1 uncultured
238	28	70.0	184	2	Q8KVI7	Q8kvi7 uncultured
239	28	70.0	184	2	Q8KVF8	Q8kvf8 uncultured
240	28	70.0	184	2	Q8KVK8	Q8kvk8 uncultured
241	28	70.0	184	2	Q8KHR7	Q8khr7 uncultured
242	28	70.0	186	11	Q8K255	Q8k255 mus musculu
243	28	70.0	193	11	P97315	P97315 mus musculu
244	28	70.0	195	16	Q9L0K7	Q9l0k7 streptomyce
245	28	70.0	208	5	Q8SR64	Q8sr64 encephalito
246	28	70.0	223	16	Q984H8	Q984h8 rhizobium l
247	28	70.0	229	12	Q7TD37	Q7td37 potato viru
248	28	70.0	229	12	Q7TD31	Q7td31 potato viru
249	28	70.0	234	16	Q7VCE1	Q7vcel prochloroco
250	28	70.0	237	16	Q7U3Z0	Q7u3z0 synechococc
251	28	70.0	239	13	Q8UUI7	Q8uui7 brachydanio
252	28	70.0	261	6	Q9XSI7	Q9xsi7 bos taurus
253	28	70.0	269	3	Q9P3V9	Q9p3v9 schizosacch
254	28	70.0	272	5	Q9XUQ7	Q9xug7 caenorhabdi
255	28	70.0	276	8	Q9TKL9	Q9tkl9 deuterocohn
256	28	70.0	280	3	Q9HGQ1	Q9hgq1 schizosacch
257	28	70.0	280	3	Q9P3E5	Q9p3e5 schizosacch
258	28	70.0	283	5	O17396	O17396 caenorhabdi
259	28	70.0	284	13	Q90Y00	Q90y00 petromyzon
260	28	70.0	291	16	Q8Y477	Q8y477 listeria mo
261	28	70.0	300	2	Q9AEC8	Q9aec8 lactobacill
262	28	70.0	300	12	Q9QB62	Q9qb62 yaba monkey
263	28	70.0	302	16	Q92DW4	Q92dw4 listeria in
264	28	70.0	302	16	Q8Y948	Q8y948 listeria mo
265	28	70.0	307	5	Q9W0H3	Q9w0h3 drosophila
266	28	70.0	311	16	Q7VJR5	Q7vjr5 helicobacte
267	28	70.0	316	13	Q805F1	Q805f1 pagrus majo
268	28	70.0	324	16	Q8A096	Q8a096 bacteroides
269	28	70.0	326	16	Q8FQ79	Q8fq79 corynebacte
270	28	70.0	332	10	Q9SN75	Q9sn75 arabidopsis
271	28	70.0	333	4	Q96A95	Q96a95 homo sapien
272	28	70.0	334	5	Q19185	Q19185 caenorhabdi
273	28	70.0	339	16	Q9PBL7	Q9pbl7 xylella fas
274	28	70.0	340	2	Q9AGN6	Q9agn6 clostridium
275	28	70.0	340	16	Q932I6	Q932i6 staphylococ
276	28	70.0	358	16	Q8YMH6	Q8ymh6 anabaena sp
277	28	70.0	369	4	Q9UPZ2	Q9upz2 homo sapien
278	28	70.0	369	5	Q9Y021	Q9y021 trypanosoma
279	28	70.0	375	10	Q9SMF0	Q9smf0 hordeum vul
280	28	70.0	395	5	Q95WD9	Q95wd9 phlebotomus
281	28	70.0	400	16	Q8DKI0	Q8dki0 synechococc
282	28	70.0	411	8	Q9MJ66	Q9mj66 physarum po
283	28	70.0	420	10	Q9C6D2	Q9c6d2 arabidopsis
284	28	70.0	423	4	Q8IUP7	Q8iup7 homo sapien
285	28	70.0	438	2	O06056	O06056 mycobacteri

286	28	70.0	438	13	Q90ZM7	Q90zm7 petromyzon
287	28	70.0	440	5	Q9Y0T4	Q9y0t4 polysphondy
288	28	70.0	443	2	Q8RTV3	Q8rtv3 uncultured
289	28	70.0	443	16	Q9PN09	Q9pn09 campylobact
290	28	70.0	444	5	O15813	O15813 dictyosteli
291	28	70.0	447	16	Q82U80	Q82u80 nitrosomona
292	28	70.0	449	5	Q9Y0T3	Q9y0t3 polysphondy
293	28	70.0	466	10	O80785	O80785 arabidopsis
294	28	70.0	470	2	Q55999	Q55999 sphingomona
295	28	70.0	470	16	Q8YFJ3	Q8y fj3 brucella me
296	28	70.0	470	16	Q8G2B8	Q8g2b8 brucella su
297	28	70.0	471	10	Q93YN7	Q93yn7 arabidopsis
298	28	70.0	479	12	Q9QDL5	Q9qdl5 sarcochilus
299	28	70.0	482	10	Q8GX97	Q8gx97 arabidopsis
300	28	70.0	484	16	Q7UYW2	Q7uyw2 rhodopirell
301	28	70.0	485	10	Q7Y0W7	Q7y0w7 lupinus alb
302	28	70.0	486	2	Q9RFC6	Q9rfc6 rhodobacter
303	28	70.0	493	16	Q8EJK4	Q8ejk4 shewanella
304	28	70.0	501	2	Q93SM4	Q93sm4 staphylococ
305	28	70.0	501	16	Q99WY9	Q99wy9 staphylococ
306	28	70.0	501	16	Q8NYI9	Q8nyi9 staphylococ
307	28	70.0	502	5	Q9GSE6	Q9gse6 tenebrio mo
308	28	70.0	522	9	Q857I8	Q857i8 mycobacteri
309	28	70.0	529	10	Q9FMV0	Q9fmv0 arabidopsis
310	28	70.0	540	10	Q9LY83	Q9ly83 arabidopsis
311	28	70.0	541	16	Q7WPA4	Q7wpa4 bordetella
312	28	70.0	551	16	Q7UDT0	Q7udt0 shigella fl
313	28	70.0	552	16	Q8XA02	Q8xa02 escherichia
314	28	70.0	554	2	Q8RTF6	Q8rtf6 oenococcus
315	28	70.0	555	2	Q8KY42	Q8ky42 streptomyce
316	28	70.0	555	2	Q8RTF7	Q8rtf7 oenococcus
317	28	70.0	558	16	Q8FL80	Q8fl80 escherichia
318	28	70.0	558	16	Q83SP3	Q83sp3 shigella fl
319	28	70.0	559	16	O83432	O83432 treponema p
320	28	70.0	617	16	Q8RDN8	Q8rdn8 fusobacteri
321	28	70.0	621	12	O38017	O38017 salmonid he
322	28	70.0	626	5	Q95ZH8	Q95zh8 cupiennius
323	28	70.0	626	5	Q95ZH9	Q95zh9 cupiennius
324	28	70.0	626	5	Q95ZH5	Q95zh5 cupiennius
325	28	70.0	626	5	Q95ZH3	Q95zh3 cupiennius
326	28	70.0	626	5	Q95ZH4	Q95zh4 cupiennius
327	28	70.0	626	5	Q95ZH7	Q95zh7 cupiennius
328	28	70.0	626	5	Q95ZH2	Q95zh2 cupiennius
329	28	70.0	626	5	Q86N90	Q86n90 nephila ina
330	28	70.0	626	5	Q86N89	Q86n89 nephila ina
331	28	70.0	629	5	Q9U1Q3	Q9u1q3 caenorhabdi
332	28	70.0	634	5	Q95ZI0	Q95zi0 cupiennius
333	28	70.0	684	5	Q9GYW1	Q9gyw1 aedes aegyp
334	28	70.0	684	5	Q9N2P8	Q9n2p8 armigeres s
335	28	70.0	685	5	Q9GYW2	Q9gyw2 aedes aegyp
336	28	70.0	690	16	Q83IB2	Q83ib2 tropheryma
337	28	70.0	694	13	Q8UUR9	Q8uur9 brachydanio
338	28	70.0	698	16	Q83GW3	Q83gw3 tropheryma
339	28	70.0	717	13	Q8AW86	Q8aw86 pagrus majo
340	28	70.0	719	5	Q9N8Z8	Q9n8z8 trypanosoma
341	28	70.0	721	12	Q9YYQ1	Q9yyq1 avian adeno
342	28	70.0	728	10	Q8S0A6	Q8s0a6 oryza sativ

343	28	70.0	743	13	Q9YGE7	Q9yge7 oncorhynch
344	28	70.0	745	10	Q7XZG7	Q7xzg7 oryza sativ
345	28	70.0	751	5	Q9U5K5	Q9u5k5 spodoptera
346	28	70.0	760	10	Q8VZU7	Q8vzu7 arabidopsis
347	28	70.0	809	5	Q9N3K6	Q9n3k6 caenorhabdi
348	28	70.0	812	9	Q858B5	Q858b5 enterobacte
349	28	70.0	826	4	Q8IY15	Q8iy15 homo sapien
350	28	70.0	861	2	Q8GID5	Q8gid5 mycoplasma
351	28	70.0	876	16	Q8XIV2	Q8xiv2 clostridium
352	28	70.0	904	5	Q86AQ1	Q86aq1 dictyosteli
353	28	70.0	963	13	Q7T3T7	Q7t3t7 oreochromis
354	28	70.0	975	3	Q877C2	Q877c2 candida tro
355	28	70.0	1011	4	O95705	O95705 homo sapien
356	28	70.0	1031	16	Q9I5Q2	Q9i5q2 pseudomonas
357	28	70.0	1073	12	Q997I4	Q997i4 bovine aden
358	28	70.0	1074	16	Q88TG3	Q88tg3 lactobacill
359	28	70.0	1255	2	Q8VUK0	Q8vuk0 paracoccus
360	28	70.0	1308	5	Q8T6H2	Q8t6h2 dictyosteli
361	28	70.0	1375	5	Q9XTF8	Q9xtf8 caenorhabdi
362	28	70.0	1450	11	O54728	O54728 rattus norv
363	28	70.0	1471	11	Q80U29	Q80u29 mus musculu
364	28	70.0	1488	4	Q9HCM3	Q9hcm3 homo sapien
365	28	70.0	1628	16	Q8DTJ7	Q8dtj7 streptococc
366	28	70.0	1957	4	O15064	O15064 homo sapien
367	28	70.0	2134	5	Q8I5M6	Q8i5m6 plasmodium
368	28	70.0	2400	5	Q8I3U7	Q8i3u7 plasmodium
369	28	70.0	2643	5	O01552	O01552 caenorhabdi
370	28	70.0	3897	12	Q98Y26	Q98y26 classical s
371	28	70.0	3898	12	Q68872	Q68872 classical s
372	28	70.0	3898	12	Q9YS30	Q9ys30 classical s
373	28	70.0	3898	12	Q991S8	Q991s8 classical s
374	28	70.0	3898	12	O92366	O92366 classical s
375	28	70.0	3898	12	O92364	O92364 classical s
376	28	70.0	3898	12	Q68965	Q68965 classical s
377	28	70.0	3898	12	Q99BK1	Q99bk1 classical s
378	28	70.0	3898	12	Q68964	Q68964 classical s
379	28	70.0	3898	12	Q68534	Q68534 classical s
380	28	70.0	3898	12	Q8JPF7	Q8jpf7 classical s
381	28	70.0	3898	12	Q68535	Q68535 classical s
382	28	70.0	3898	12	Q96891	Q96891 classical s
383	28	70.0	3898	12	O09710	O09710 pestivirus
384	28	70.0	3898	12	Q68871	Q68871 classical s
385	28	70.0	3898	12	Q80BB0	Q80bb0 classical s
386	28	70.0	4496	5	Q9VDG0	Q9vdg0 drosophila
387	28	70.0	5251	5	Q8IID4	Q8iid4 plasmodium
388	27	67.5	50	16	Q97SP9	Q97sp9 streptococc
389	27	67.5	58	17	Q8ZV06	Q8zv06 pyrobaculum
390	27	67.5	69	16	Q7VH03	Q7vh03 helicobacte
391	27	67.5	72	16	Q8FJ80	Q8fj80 escherichia
392	27	67.5	76	4	Q9H3A8	Q9h3a8 homo sapien
393	27	67.5	82	16	Q8KB13	Q8kb13 chlorobium
394	27	67.5	88	16	Q7UR07	Q7ur07 rhodopirell
395	27	67.5	92	16	Q9HWW8	Q9hww8 pseudomonas
396	27	67.5	104	16	Q9WZN2	Q9wzn2 thermotoga
397	27	67.5	105	9	Q7Y598	Q7y598 bacterioph
398	27	67.5	109	12	Q9DR78	Q9dr78 bovine coro
399	27	67.5	109	12	Q84732	Q84732 porcine hem

400	27	67.5	109	12	Q9QAQ5	Q9qaq5 bovine coro
401	27	67.5	109	12	Q8JSP6	Q8jsp6 porcine hem
402	27	67.5	109	12	Q99H66	Q99h66 bovine coro
403	27	67.5	109	12	Q8V434	Q8v434 bovine coro
404	27	67.5	109	12	Q9PY31	Q9py31 bovine coro
405	27	67.5	109	12	Q89792	Q89792 porcine hem
406	27	67.5	109	12	Q9QAR2	Q9qar2 bovine coro
407	27	67.5	111	16	Q9ACQ0	Q9acq0 streptomyce
408	27	67.5	117	9	Q9T0Y0	Q9t0y0 lactobacill
409	27	67.5	119	16	Q8D9V8	Q8d9v8 vibrio vuln
410	27	67.5	128	2	Q8GHA1	Q8gha1 actinobacil
411	27	67.5	133	16	Q9ADJ5	Q9adj5 streptomyce
412	27	67.5	136	16	Q987H0	Q987h0 rhizobium l
413	27	67.5	142	5	Q21371	Q21371 caenorhabdi
414	27	67.5	143	2	Q8GM58	Q8gm58 mycoplasma
415	27	67.5	148	5	Q9GRP2	Q9grp2 leishmania
416	27	67.5	149	17	O51955	O51955 halobacteri
417	27	67.5	151	16	Q8E234	Q8e234 streptococc
418	27	67.5	153	16	Q99TR0	Q99tr0 staphylococ
419	27	67.5	153	16	Q8E7J0	Q8e7j0 streptococc
420	27	67.5	159	12	Q91GH2	Q91gh2 epiphyas po
421	27	67.5	159	12	Q7TLR2	Q7tlr2 choristoneu
422	27	67.5	159	13	Q8AYG7	Q8ayg7 gallus gall
423	27	67.5	162	8	Q95AX5	Q95ax5 dasypogon h
424	27	67.5	163	16	Q9CE20	Q9ce20 lactococcus
425	27	67.5	164	2	Q9X9R9	Q9x9r9 streptococc
426	27	67.5	167	16	Q9PM12	Q9pml2 campylobact
427	27	67.5	174	16	Q92CD6	Q92cd6 listeria in
428	27	67.5	176	2	Q8RT79	Q8rt79 listeria mo
429	27	67.5	177	5	Q20070	Q20070 caenorhabdi
430	27	67.5	182	16	Q835H4	Q835h4 enterococcu
431	27	67.5	184	2	Q8KVL0	Q8kvl0 uncultured
432	27	67.5	184	16	Q813S8	Q813s8 bacillus ce
433	27	67.5	186	16	Q92BN0	Q92bn0 listeria in
434	27	67.5	186	16	Q8Y739	Q8y739 listeria mo
435	27	67.5	188	16	Q8F442	Q8f442 leptospira
436	27	67.5	189	5	Q8IP62	Q8ip62 drosophila
437	27	67.5	190	17	Q9HSH9	Q9hsh9 halobacteri
438	27	67.5	203	4	Q9H3I4	Q9h3i4 homo sapien
439	27	67.5	203	16	Q8NN67	Q8nn67 corynebacte
440	27	67.5	210	16	Q931V6	Q931v6 staphylococ
441	27	67.5	212	10	Q9M313	Q9m313 arabidopsis
442	27	67.5	216	9	O64061	O64061 bacterioph
443	27	67.5	216	16	O31960	O31960 bacillus su
444	27	67.5	223	16	Q8RDM7	Q8rdm7 fusobacteri
445	27	67.5	229	5	Q86FN0	Q86fn0 caenorhabdi
446	27	67.5	230	16	Q9PHW2	Q9phw2 campylobact
447	27	67.5	235	1	P71164	P71164 desulfuroco
448	27	67.5	236	16	Q9PKB7	Q9pkb7 chlamydia m
449	27	67.5	239	16	Q9RL46	Q9rl46 streptomyce
450	27	67.5	239	16	Q7V8E0	Q7v8e0 prochloroco
451	27	67.5	241	10	Q9LJA4	Q9lja4 arabidopsis
452	27	67.5	245	16	Q9CKD2	Q9ckd2 pasteurella
453	27	67.5	245	16	Q98AZ2	Q98az2 rhizobium l
454	27	67.5	246	16	Q83L46	Q83l46 shigella fl
455	27	67.5	248	16	Q8ZG63	Q8zg63 yersinia pe
456	27	67.5	258	16	Q8FBI6	Q8fbi6 escherichia

457	27	67.5	258	16	Q8CT67	Q8ct67	staphylococ
458	27	67.5	258	16	Q83PG7	Q83pg7	shigella fl
459	27	67.5	260	10	Q7XHJ8	Q7xhj8	lycopersico
460	27	67.5	263	10	O65102	O65102	actinidia d
461	27	67.5	266	5	Q8SV83	Q8sv83	encephalito
462	27	67.5	268	10	Q9FP68	Q9fp68	oryza sativ
463	27	67.5	272	4	Q7Z2S1	Q7z2s1	homo sapien
464	27	67.5	273	16	Q8YMB8	Q8ymb8	anabaena sp
465	27	67.5	281	16	Q8EUJ6	Q8euj6	mycoplasma
466	27	67.5	281	16	Q83EA8	Q83ea8	coxiella bu
467	27	67.5	281	17	O29061	O29061	archaeoglob
468	27	67.5	283	11	Q8CBI4	Q8cbi4	mus musculu
469	27	67.5	284	17	Q971G0	Q971g0	sulfolobus
470	27	67.5	286	2	Q8KMW1	Q8kmw1	escherichia
471	27	67.5	286	10	Q9SJI5	Q9sji5	arabidopsis
472	27	67.5	286	11	Q7TMW7	Q7tmw7	mus musculu
473	27	67.5	289	16	Q8FG37	Q8fg37	escherichia
474	27	67.5	289	16	Q9Z6I7	Q9z6i7	chlamydia p
475	27	67.5	293	10	Q9XIW1	Q9xiw1	arabidopsis
476	27	67.5	293	10	Q41614	Q41614	tropaeolum
477	27	67.5	293	10	P93046	P93046	arabidopsis
478	27	67.5	293	10	Q9LXH6	Q9lxh6	arabidopsis
479	27	67.5	295	10	O65734	O65734	cicer ariet
480	27	67.5	297	10	Q93XM0	Q93xm0	carica papa
481	27	67.5	300	5	Q9XU61	Q9xu61	caenorhabdi
482	27	67.5	301	17	Q8U168	Q8u168	pyrococcus
483	27	67.5	302	16	Q81FD8	Q81fd8	bacillus ce
484	27	67.5	308	10	Q8LD18	Q8ld18	arabidopsis
485	27	67.5	309	16	Q7UV34	Q7uv34	rhodopirell
486	27	67.5	310	3	Q99190	Q99190	saccharomyc
487	27	67.5	313	16	O25337	O25337	helicobacte
488	27	67.5	313	16	Q9ZLM0	Q9zlm0	helicobacte
489	27	67.5	313	16	Q8DHK0	Q8dhk0	synechococc
490	27	67.5	314	16	Q9K834	Q9k834	bacillus ha
491	27	67.5	317	13	Q7T180	Q7t180	gallus gall
492	27	67.5	318	16	Q8D5G7	Q8d5g7	vibrio vuln
493	27	67.5	318	17	Q96Z66	Q96z66	sulfolobus
494	27	67.5	319	16	Q8ENW8	Q8enw8	oceanobacil
495	27	67.5	323	16	Q8YDY7	Q8ydy7	brucella me
496	27	67.5	324	16	Q92J31	Q92j31	rickettsia
497	27	67.5	328	5	Q9N3W0	Q9n3w0	caenorhabdi
498	27	67.5	328	16	Q9RPS4	Q9rps4	enterococcu
499	27	67.5	331	4	Q8N9S3	Q8n9s3	homo sapien
500	27	67.5	336	13	Q8AVS1	Q8avs1	xenopus lae
501	27	67.5	337	16	Q8ZL16	Q8zl16	salmonella
502	27	67.5	337	16	Q8Z2K8	Q8z2k8	salmonella
503	27	67.5	337	16	Q83SU3	Q83su3	salmonella
504	27	67.5	338	13	Q7SXN3	Q7sxn3	brachydanio
505	27	67.5	339	10	Q9SFZ7	Q9sfz7	arabidopsis
506	27	67.5	345	16	Q97RN7	Q97rn7	streptococc
507	27	67.5	345	16	Q8DQG8	Q8dqq8	streptococc
508	27	67.5	349	5	Q9XU59	Q9xu59	caenorhabdi
509	27	67.5	354	5	Q9V9Q4	Q9v9q4	drosophila
510	27	67.5	355	16	Q87XV7	Q87xv7	pseudomonas
511	27	67.5	358	4	O75769	O75769	homo sapien
512	27	67.5	361	16	Q92TH6	Q92th6	rhizobium m
513	27	67.5	365	12	Q91TQ9	Q91tq9	tupaia herp

514	27	67.5	366	10	Q9M7D5	Q9m7d5 ipomoea bat
515	27	67.5	366	10	Q9AUC5	Q9auc5 ipomoea bat
516	27	67.5	368	10	Q9M7D6	Q9m7d6 ipomoea bat
517	27	67.5	368	10	Q84RM9	Q84rm9 ipomoea bat
518	27	67.5	368	10	Q84RM8	Q84rm8 ipomoea bat
519	27	67.5	371	16	Q8U7G2	Q8u7g2 agrobacteri
520	27	67.5	374	16	Q927X3	Q927x3 listeria in
521	27	67.5	383	16	Q8FXK8	Q8fxk8 brucella su
522	27	67.5	384	8	Q8HDG6	Q8hdg6 pteridomona
523	27	67.5	385	16	Q9AA68	Q9aa68 caulobacter
524	27	67.5	387	8	Q7YMT9	Q7ymt9 dictyocha f
525	27	67.5	387	16	Q81FH5	Q81fh5 bacillus ce
526	27	67.5	405	8	Q8MAA4	Q8maa4 rhodella vi
527	27	67.5	405	10	Q9SXC4	Q9sxc4 arabidopsis
528	27	67.5	410	17	Q979W6	Q979w6 thermoplasm
529	27	67.5	413	5	Q9XZR6	Q9xZR6 drosophila
530	27	67.5	415	10	Q8RWT4	Q8rwt4 arabidopsis
531	27	67.5	422	16	Q97SC5	Q97sc5 streptococc
532	27	67.5	428	5	Q8SXA2	Q8sxa2 drosophila
533	27	67.5	430	16	Q8FW04	Q8fw04 brucella su
534	27	67.5	433	10	Q9SZS1	Q9szs1 arabidopsis
535	27	67.5	434	5	Q23217	Q23217 caenorhabdi
536	27	67.5	436	8	Q8M9J3	Q8m9j3 ilex pubesc
537	27	67.5	437	13	Q9DEQ5	Q9deq5 gallus gall
538	27	67.5	440	6	Q95K33	Q95k33 macaca fasc
539	27	67.5	441	16	Q89RS5	Q89rs5 bradyrhizob
540	27	67.5	444	11	Q9CV28	Q9cv28 mus musculu
541	27	67.5	445	3	Q94346	Q94346 schizosacch
542	27	67.5	445	4	Q9H8M7	Q9h8m7 homo sapien
543	27	67.5	448	16	Q87M09	Q87m09 vibrio para
544	27	67.5	449	16	Q81HX4	Q81hx4 bacillus ce
545	27	67.5	451	10	Q9FNU6	Q9fnu6 populus x c
546	27	67.5	454	10	Q9FUE7	Q9fue7 populus del
547	27	67.5	454	10	Q9FUE9	Q9fue9 populus tri
548	27	67.5	454	10	Q9FUE8	Q9fue8 populus tri
549	27	67.5	455	4	Q96S13	Q96s13 homo sapien
550	27	67.5	455	16	Q8CZE5	Q8cze5 oceanobacil
551	27	67.5	455	17	Q50563	Q50563 methanobact
552	27	67.5	457	8	Q8HDG7	Q8hdg7 pedinella s
553	27	67.5	459	5	Q20436	Q20436 caenorhabdi
554	27	67.5	461	16	Q7VHF0	Q7vhf0 helicobacte
555	27	67.5	462	2	Q9RA87	Q9ra87 sphingomona
556	27	67.5	462	2	Q9RP80	Q9rp80 pseudomonas
557	27	67.5	468	3	Q9P8W6	Q9p8w6 candida alb
558	27	67.5	473	5	Q9VN12	Q9vn12 drosophila
559	27	67.5	482	10	Q9FPT4	Q9fpt4 arabidopsis
560	27	67.5	482	10	Q949Y0	Q949y0 arabidopsis
561	27	67.5	483	12	Q82172	Q82172 leporid her
562	27	67.5	488	10	Q8LPN8	Q8lpn8 arabidopsis
563	27	67.5	489	16	Q8XPS9	Q8xps9 ralstonia s
564	27	67.5	490	10	Q9C8H9	Q9c8h9 arabidopsis
565	27	67.5	500	16	Q8XCA2	Q8xca2 escherichia
566	27	67.5	500	16	Q8FAC8	Q8fac8 escherichia
567	27	67.5	502	13	Q800I1	Q800il cynops pyrr
568	27	67.5	503	2	Q8GJK0	Q8gjk0 meiothermus
569	27	67.5	504	2	Q9ZIP5	Q9zip5 rhizobium l
570	27	67.5	504	16	Q7UAP9	Q7uap9 shigella fl

571	27	67.5	512	12	O56677	O56677	cocal virus
572	27	67.5	513	16	Q9ZDP5	Q9zdp5	rickettsia
573	27	67.5	513	16	Q8ZJE1	Q8zjel	yersinia pe
574	27	67.5	515	16	Q985L5	Q985l5	rhizobium l
575	27	67.5	516	16	Q9ABY0	Q9aby0	caulobacter
576	27	67.5	516	16	Q83IM7	Q83im7	shigella fl
577	27	67.5	518	5	P98186	P98186	boophilus m
578	27	67.5	520	16	Q92RC7	Q92rc7	rhizobium m
579	27	67.5	521	16	Q92DS3	Q92ds3	listeria in
580	27	67.5	522	10	Q9LPY0	Q9lpy0	arabidopsis
581	27	67.5	524	16	Q82FS7	Q82fs7	streptomyce
582	27	67.5	527	5	Q9N4T6	Q9n4t6	caenorhabdi
583	27	67.5	529	17	Q97B77	Q97b77	thermoplasm
584	27	67.5	533	16	O06173	O06173	mycobacteri
585	27	67.5	533	16	Q7TYF0	Q7tyf0	mycobacteri
586	27	67.5	537	4	Q96AA0	Q96aa0	homo sapien
587	27	67.5	539	16	Q89QR2	Q89qr2	bradyrhizob
588	27	67.5	541	10	Q851K4	Q851k4	oryza sativ
589	27	67.5	549	8	Q9G910	Q9g910	ochromonas
590	27	67.5	553	16	Q8F932	Q8f932	leptospira
591	27	67.5	564	16	Q89R39	Q89r39	bradyrhizob
592	27	67.5	569	2	Q93CY6	Q93cy6	lactobacill
593	27	67.5	570	10	Q9LU36	Q9lu36	arabidopsis
594	27	67.5	570	10	Q84P22	Q84p22	arabidopsis
595	27	67.5	571	16	Q8XHF8	Q8xhf8	clostridium
596	27	67.5	576	16	O06137	O06137	mycobacteri
597	27	67.5	576	16	Q7TZV7	Q7tzv7	mycobacteri
598	27	67.5	597	10	O04659	O04659	arabidopsis
599	27	67.5	598	10	O82747	O82747	arabidopsis
600	27	67.5	604	5	Q9U1E2	Q9ule2	leishmania
601	27	67.5	606	11	Q91VU5	Q91vu5	mus musculu
602	27	67.5	612	3	Q870R2	Q870r2	neurospora
603	27	67.5	613	5	Q27334	Q27334	onchocerca
604	27	67.5	616	10	Q9LV10	Q9lv10	arabidopsis
605	27	67.5	628	5	Q8IPM4	Q8ipm4	drosophila
606	27	67.5	629	16	Q7V3R9	Q7v3r9	prochloroco
607	27	67.5	631	16	Q9KE29	Q9ke29	bacillus ha
608	27	67.5	631	16	Q8EN36	Q8en36	oceanobacil
609	27	67.5	631	16	Q81YT2	Q81yt2	bacillus an
610	27	67.5	631	16	Q81I69	Q81i69	bacillus ce
611	27	67.5	634	8	Q9MRP1	Q9mrp1	glomeropitc
612	27	67.5	640	10	Q8W0B9	Q8w0b9	oryza sativ
613	27	67.5	650	17	Q8TSE0	Q8tse0	methanosarc
614	27	67.5	651	12	Q9PYY4	Q9pyy4	xestia c-ni
615	27	67.5	651	17	Q8THF4	Q8thf4	methanosarc
616	27	67.5	659	2	Q8RP47	Q8rp47	azospirillu
617	27	67.5	660	8	Q7YMD7	Q7ymd7	spatholirio
618	27	67.5	666	10	Q8H0J9	Q8h0j9	triticum ur
619	27	67.5	691	8	Q9MRP6	Q9mrp6	catopsis wa
620	27	67.5	695	16	Q83DR9	Q83dr9	coxiella bu
621	27	67.5	696	16	Q8A0L4	Q8a0l4	bacteroides
622	27	67.5	698	5	Q9U5K4	Q9u5k4	spodoptera
623	27	67.5	699	10	Q93ZQ5	Q93zq5	arabidopsis
624	27	67.5	703	5	Q9GQ56	Q9gq56	corcyra cep
625	27	67.5	703	10	Q9ZSB4	Q9zsb4	arabidopsis
626	27	67.5	712	5	O16163	O16163	drosophila
627	27	67.5	715	4	Q7Z5X1	Q7z5x1	homo sapien

628	27	67.5	749	12	Q7T4K3	Q7t4k3 pea stem ne
629	27	67.5	779	4	Q9H3P4	Q9h3p4 homo sapien
630	27	67.5	779	16	Q9ZJK9	Q9zjk9 helicobacte
631	27	67.5	783	16	Q9A842	Q9a842 caulobacter
632	27	67.5	784	10	Q9FVA3	Q9fva3 zea mays (m
633	27	67.5	796	3	Q9P937	Q9p937 candida alb
634	27	67.5	800	3	Q96V31	Q96v31 candida alb
635	27	67.5	840	2	Q8KLI9	Q8kli9 thermomonos
636	27	67.5	842	16	Q8PN87	Q8pn87 xanthomonas
637	27	67.5	847	16	Q8YXR5	Q8yxr5 anabaena sp
638	27	67.5	848	2	Q84HQ5	Q84hq5 helicobacte
639	27	67.5	861	16	Q8PBN1	Q8pbn1 xanthomonas
640	27	67.5	879	2	O30579	O30579 bacillus su
641	27	67.5	879	16	O31504	O31504 bacillus su
642	27	67.5	899	3	Q8TFN6	Q8tfn6 colletotric
643	27	67.5	904	2	O50187	O50187 mycoplasma
644	27	67.5	915	16	Q7VCW0	Q7vcw0 prochloroco
645	27	67.5	930	16	Q8E6N1	Q8e6n1 streptococc
646	27	67.5	930	16	Q8E177	Q8e177 streptococc
647	27	67.5	930	16	Q8DVD3	Q8dvd3 streptococc
648	27	67.5	930	16	Q8DNW4	Q8dnw4 streptococc
649	27	67.5	932	16	Q88V89	Q88v89 lactobacill
650	27	67.5	933	16	Q99YW3	Q99yw3 streptococc
651	27	67.5	933	16	Q8P068	Q8p068 streptococc
652	27	67.5	933	16	Q8K6S2	Q8k6s2 streptococc
653	27	67.5	937	4	Q96FT5	Q96ft5 homo sapien
654	27	67.5	949	16	Q9KU47	Q9ku47 vibrio chol
655	27	67.5	951	13	Q7ZUF4	Q7zuf4 brachydanio
656	27	67.5	955	16	Q9X1N2	Q9xln2 thermotoga
657	27	67.5	956	11	Q99K64	Q99k64 mus musculu
658	27	67.5	956	11	Q8R542	Q8r542 mus musculu
659	27	67.5	956	16	Q8P6H7	Q8p6h7 xanthomonas
660	27	67.5	995	3	Q9UUE8	Q9uue8 schizosacch
661	27	67.5	996	16	Q97NR7	Q97nr7 streptococc
662	27	67.5	996	16	Q8DNF6	Q8dnf6 streptococc
663	27	67.5	1014	10	Q84KC5	Q84kc5 hordeum vul
664	27	67.5	1016	4	Q8NDE6	Q8nde6 homo sapien
665	27	67.5	1019	4	O94947	O94947 homo sapien
666	27	67.5	1024	11	Q9ESJ5	Q9esj5 mus musculu
667	27	67.5	1050	8	Q7YN59	Q7yn59 eimeria ten
668	27	67.5	1056	10	Q7XV85	Q7xv85 oryza sativ
669	27	67.5	1059	10	Q8LI36	Q8li36 oryza sativ
670	27	67.5	1124	16	Q9CM76	Q9cm76 pasteurella
671	27	67.5	1168	10	Q9LE81	Q9le81 arabidopsis
672	27	67.5	1196	10	Q9SG02	Q9sg02 arabidopsis
673	27	67.5	1196	10	Q9LKP0	Q9lkp0 arabidopsis
674	27	67.5	1218	10	Q8LHH9	Q8lhh9 oryza sativ
675	27	67.5	1235	3	Q8TFI5	Q8tfi5 saccharomyc
676	27	67.5	1297	3	Q9Y817	Q9y817 schizosacch
677	27	67.5	1451	5	O61779	O61779 caenorhabdi
678	27	67.5	1457	10	Q8LPU6	Q8lpu6 zea mays (m
679	27	67.5	1458	5	Q9VX48	Q9vx48 drosophila
680	27	67.5	1470	12	Q8JYK1	Q8jyk1 fiji diseas
681	27	67.5	1512	11	Q8K3K1	Q8k3k1 rattus norv
682	27	67.5	1525	10	O65343	O65343 zea mays (m
683	27	67.5	1554	5	Q8IJY4	Q8ijy4 plasmodium
684	27	67.5	1557	11	Q811F2	Q811f2 mus musculu



685	27	67.5	1620	5	O45000	O45000 caenorhabdi
686	27	67.5	1725	6	Q867A1	Q867a1 canis famil
687	27	67.5	1725	11	P70570	P70570 rattus norv
688	27	67.5	1740	5	Q86B45	Q86b45 drosophila
689	27	67.5	1743	4	Q9BYP7	Q9byp7 homo sapien
690	27	67.5	1800	4	Q8TCX6	Q8tcx6 homo sapien
691	27	67.5	1806	4	Q96TG0	Q96tg0 homo sapien
692	27	67.5	2043	5	Q8WQ87	Q8wq87 chironomus
693	27	67.5	2184	12	Q9DXZ1	Q9dxz1 canine dist
694	27	67.5	2184	12	Q9QPQ7	Q9qpq7 canine dist
695	27	67.5	2184	12	P87575	P87575 canine dist
696	27	67.5	2184	12	Q91KN1	Q91kn1 canine dist
697	27	67.5	2219	4	Q9C0A3	Q9c0a3 homo sapien
698	27	67.5	2296	4	Q9Y3S1	Q9y3s1 homo sapien
699	27	67.5	2367	2	Q9EXR0	Q9exr0 clostridium
700	27	67.5	3068	12	Q8Jrq5	Q8jrq5 pepper mott
701	27	67.5	3364	5	Q8IM60	Q8im60 plasmodium
702	27	67.5	3610	5	Q968T7	Q968t7 plasmodium
703	27	67.5	3620	5	Q968T6	Q968t6 plasmodium
704	27	67.5	3628	5	Q968Y0	Q968y0 plasmodium
705	27	67.5	3704	5	Q8IKY8	Q8iky8 plasmodium
706	27	67.5	4564	5	O77075	O77075 drosophila
707	27	67.5	5687	5	Q8IJH4	Q8ijh4 plasmodium
708	26	65.0	15	5	Q9TWF5	Q9twf5 artemia (br
709	26	65.0	25	4	Q8WYP6	Q8wyp6 homo sapien
710	26	65.0	32	16	O25168	O25168 helicobacte
711	26	65.0	60	5	Q9BHZ8	Q9bhz8 globodera r
712	26	65.0	68	16	Q7UYG6	Q7uyg6 rhodopirell
713	26	65.0	75	12	Q90160	Q90160 bombyx mori
714	26	65.0	77	16	Q8CNE4	Q8cne4 staphylococ
715	26	65.0	80	10	Q9SJ65	Q9sj65 arabidopsis
716	26	65.0	85	5	Q19641	Q19641 caenorhabdi
717	26	65.0	85	16	Q8YGA7	Q8yga7 brucella me
718	26	65.0	85	16	Q8G1L7	Q8g1l7 brucella su
719	26	65.0	88	1	Q9HHC5	Q9hhc5 halobacteri
720	26	65.0	95	2	Q8VQR5	Q8vqr5 escherichia
721	26	65.0	100	16	Q92D77	Q92d77 listeria in
722	26	65.0	101	8	Q952Q4	Q952q4 euphorbia p
723	26	65.0	101	16	Q894W9	Q894w9 clostridium
724	26	65.0	101	16	Q835U9	Q835u9 enterococcu
725	26	65.0	104	16	Q8YTV9	Q8ytv9 anabaena sp
726	26	65.0	111	10	Q8GSD4	Q8gsd4 oryza sativ
727	26	65.0	115	16	Q8XGD8	Q8xgd8 salmonella
728	26	65.0	116	16	Q9CJC8	Q9cjc8 lactococcus
729	26	65.0	122	6	O77629	O77629 sus scrofa
730	26	65.0	122	16	Q9K5Q1	Q9k5q1 bacillus ha
731	26	65.0	124	10	Q9LGZ5	Q9lgz5 oryza sativ
732	26	65.0	132	10	Q8S812	Q8s812 oryza sativ
733	26	65.0	132	16	Q9ZBJ6	Q9zbj6 streptomyce
734	26	65.0	132	16	Q7U640	Q7u640 synechococc
735	26	65.0	133	2	Q9FAY8	Q9fay8 streptococc
736	26	65.0	136	6	Q95295	Q95295 sus scrofa
737	26	65.0	136	6	Q9TS88	Q9ts88 bos taurus
738	26	65.0	137	16	Q9PPJ2	Q9ppj2 campylobact
739	26	65.0	139	17	Q9HHV8	Q9hhv8 halobacteri
740	26	65.0	140	16	Q8E720	Q8e720 streptococc
741	26	65.0	140	16	Q8E1K3	Q8elk3 streptococc

742	26	65.0	144	16	Q9X244	Q9x244 thermotoga
743	26	65.0	148	13	Q8AYG6	Q8ayg6 gallus gall
744	26	65.0	151	11	Q8CB29	Q8cb29 mus musculu
745	26	65.0	152	17	Q8TT66	Q8tt66 methanosarc
746	26	65.0	154	8	Q9XM12	Q9xml2 cervus elap
747	26	65.0	154	13	Q9DFF0	Q9dff0 oncorhynchu
748	26	65.0	155	16	Q8NMS6	Q8nms6 corynebacte
749	26	65.0	156	16	Q8X949	Q8x949 escherichia
750	26	65.0	156	16	Q8UAD1	Q8uad1 agrobacteri
751	26	65.0	156	16	Q7UDR8	Q7udr8 shigella fl
752	26	65.0	158	10	Q8LC62	Q8lc62 arabidopsis
753	26	65.0	158	10	Q9SYL3	Q9syl3 arabidopsis
754	26	65.0	158	11	Q8C8D9	Q8c8d9 mus musculu
755	26	65.0	159	5	O16738	O16738 caenorhabdi
756	26	65.0	162	8	O47208	O47208 anigozantho
757	26	65.0	162	10	Q8S0K1	Q8s0k1 oryza sativ
758	26	65.0	162	16	Q9Z6W2	Q9z6w2 chlamydia p
759	26	65.0	163	10	Q8GWP6	Q8gwp6 arabidopsis
760	26	65.0	164	16	Q836E9	Q836e9 enterococcu
761	26	65.0	167	5	Q9N4B5	Q9n4b5 caenorhabdi
762	26	65.0	168	17	Q8ZWK1	Q8zwk1 pyrobaculum
763	26	65.0	169	10	Q9LGC1	Q9lgc1 oryza sativ
764	26	65.0	169	10	Q946K1	Q946k1 uncultured
765	26	65.0	170	8	Q85WI4	Q85wi4 uncultured
766	26	65.0	171	2	Q93DS0	Q93ds0 uncultured
767	26	65.0	171	10	Q946K9	Q946k9 uncultured
768	26	65.0	172	16	Q83SM3	Q83sm3 shigella fl
769	26	65.0	179	2	O87062	O87062 vibrio chol
770	26	65.0	179	15	Q7STF3	Q7stf3 human immun
771	26	65.0	182	10	P93190	P93190 hordeum vul
772	26	65.0	183	16	P96579	P96579 bacillus su
773	26	65.0	183	16	Q9RQ09	Q9rq09 bacteroides
774	26	65.0	184	2	Q8KUZ4	Q8kuz4 uncultured
775	26	65.0	185	16	Q89NA0	Q89na0 bradyrhizob
776	26	65.0	186	10	P93414	P93414 oryza sativ
777	26	65.0	186	16	Q8EPV6	Q8epv6 oceanobacil
778	26	65.0	190	4	Q8ND18	Q8ndl8 homo sapien
779	26	65.0	190	10	Q07525	Q07525 tropaeolum
780	26	65.0	190	15	Q7ZS86	Q7zs86 human immun
781	26	65.0	190	16	Q8FMW2	Q8fmw2 corynebacte
782	26	65.0	192	16	P72764	P72764 synechocyst
783	26	65.0	196	16	Q82IE5	Q82ie5 streptomyce
784	26	65.0	197	10	Q94F76	Q94f76 zea mays (m
785	26	65.0	199	13	Q8JHI9	Q8jhi9 brachydanio
786	26	65.0	200	15	Q9QLA1	Q9qla1 human immun
787	26	65.0	205	16	Q98C34	Q98c34 rhizobium l
788	26	65.0	205	16	Q9X006	Q9x006 thermotoga
789	26	65.0	206	13	Q803P7	Q803p7 brachydanio
790	26	65.0	208	11	Q9DCT8	Q9dct8 mus musculu
791	26	65.0	211	5	Q9N5K2	Q9n5k2 caenorhabdi
792	26	65.0	212	10	Q7XKH5	Q7xkh5 oryza sativ
793	26	65.0	214	10	Q9LZY2	Q9lzy2 arabidopsis
794	26	65.0	214	16	Q9CIJ6	Q9cij6 lactococcus
795	26	65.0	215	4	Q9UFX7	Q9ufx7 homo sapien
796	26	65.0	215	10	Q7XQR5	Q7xqr5 oryza sativ
797	26	65.0	215	16	Q87IG4	Q87ig4 vibrio para
798	26	65.0	219	13	Q7ZUX7	Q7zux7 brachydanio

799	26	65.0	221	16	Q930S1	Q930s1 rhizobium m
800	26	65.0	221	17	O58889	O58889 pyrococcus
801	26	65.0	222	10	Q7XT07	Q7xt07 oryza sativ
802	26	65.0	224	3	Q05243	Q05243 saccharomyc
803	26	65.0	224	5	Q95QS6	Q95qs6 caenorhabdi
804	26	65.0	225	16	Q89AZ8	Q89az8 buchnera ap
805	26	65.0	226	5	O17751	O17751 caenorhabdi
806	26	65.0	229	9	Q9MC82	Q9mc82 bacterioph
807	26	65.0	230	5	Q95R15	Q95r15 caenorhabdi
808	26	65.0	231	11	Q8C6Y8	Q8c6y8 mus musculu
809	26	65.0	233	13	Q9PVT0	Q9pvt0 oryzias lat
810	26	65.0	234	16	Q9X1F3	Q9x1f3 thermotoga
811	26	65.0	237	16	Q8YXN8	Q8yxn8 anabaena sp
812	26	65.0	239	9	Q8SDX2	Q8sdx2 bacterioph
813	26	65.0	242	13	Q7T1E2	Q7tle2 gallus gall
814	26	65.0	244	16	Q8D3B0	Q8d3b0 wiggleswort
815	26	65.0	244	16	Q7VR20	Q7vr20 candidatus
816	26	65.0	248	2	Q8VUF2	Q8vuf2 azoarcus ev
817	26	65.0	248	5	O62113	O62113 caenorhabdi
818	26	65.0	248	16	Q9ZCQ5	Q9zqc5 rickettsia
819	26	65.0	249	3	Q9P6I6	Q9p6i6 schizosacch
820	26	65.0	249	16	Q7UZN7	Q7uzn7 prochloroco
821	26	65.0	250	2	Q8RPW9	Q8rpw9 bacillus sp
822	26	65.0	250	16	Q8UCK3	Q8uck3 agrobacteri
823	26	65.0	251	17	Q9V2E8	Q9v2e8 pyrococcus
824	26	65.0	252	8	P92337	P92337 pleurothall
825	26	65.0	252	16	Q7V020	Q7v020 prochloroco
826	26	65.0	254	16	Q88Y29	Q88y29 lactobacill
827	26	65.0	256	9	O80070	O80070 staphylococ
828	26	65.0	256	16	Q8NVM4	Q8nvm4 staphylococ
829	26	65.0	257	8	Q9T6M2	Q9t6m2 globodera p
830	26	65.0	257	16	Q45494	Q45494 bacillus su
831	26	65.0	258	5	Q7Z1C4	Q7z1c4 plasmodium
832	26	65.0	258	16	Q8FLW6	Q8flw6 corynebacte
833	26	65.0	259	5	O17708	O17708 caenorhabdi
834	26	65.0	259	16	Q9L6M3	Q9l6m3 s sec-indep
835	26	65.0	260	16	Q9AA01	Q9aa01 caulobacter
836	26	65.0	260	16	O34516	O34516 bacillus su
837	26	65.0	262	6	Q9BGL3	Q9bgl3 ovis aries
838	26	65.0	262	17	Q8TTP6	Q8ttp6 methanosarc
839	26	65.0	264	10	Q9ZQZ0	Q9zqz0 oryza sativ
840	26	65.0	264	10	Q8L9W1	Q8l9w1 arabidopsis
841	26	65.0	264	10	O23619	O23619 arabidopsis
842	26	65.0	265	16	Q9Z9J1	Q9z9j1 bacillus ha
843	26	65.0	267	10	Q9SMP1	Q9smp1 arabidopsis
844	26	65.0	269	12	Q8JTN6	Q8jtn6 lumpy skin
845	26	65.0	269	16	Q8R6A5	Q8r6a5 fusobacteri
846	26	65.0	272	2	Q8KGL0	Q8kg10 rhizobium l
847	26	65.0	272	16	Q98AR5	Q98ar5 rhizobium l
848	26	65.0	274	5	Q8MXG1	Q8mxg1 caenorhabdi
849	26	65.0	274	5	Q9NAK3	Q9nak3 caenorhabdi
850	26	65.0	276	16	Q9AAJ9	Q9aaj9 caulobacter
851	26	65.0	277	11	Q8RLV2	Q8rlv2 mus musculu
852	26	65.0	280	5	Q8ISI9	Q8isi9 perkinsus m
853	26	65.0	282	16	Q8EWR3	Q8ewr3 mycoplasma
854	26	65.0	284	12	Q918K7	Q918k7 vaccinia vi
855	26	65.0	284	12	Q918K5	Q918k5 vaccinia vi

856	26	65.0	284	12	Q918K3	Q918k3	vaccinia vi
857	26	65.0	285	10	O48777	O48777	arabidopsis
858	26	65.0	286	16	Q98PE3	Q98pe3	mycoplasma
859	26	65.0	286	16	Q7VLI3	Q7vli3	prochloroco
860	26	65.0	289	17	Q9HIY5	Q9hiy5	thermoplasm
861	26	65.0	292	17	Q9Y9R9	Q9y9r9	aeropyrum p
862	26	65.0	293	10	Q38696	Q38696	actinidia d
863	26	65.0	293	12	Q65412	Q65412	bean common
864	26	65.0	294	16	Q7WEG6	Q7weg6	bordetella
865	26	65.0	294	16	Q7W349	Q7w349	bordetella
866	26	65.0	294	16	Q7W033	Q7w033	bordetella
867	26	65.0	295	4	Q9H746	Q9h746	homo sapien
868	26	65.0	295	10	Q07524	Q07524	tropaeolum
869	26	65.0	297	17	Q8PSL3	Q8psl3	methanosarc
870	26	65.0	300	5	Q23614	Q23614	caenorhabdi
871	26	65.0	300	16	Q9A8B0	Q9a8b0	caulobacter
872	26	65.0	300	16	Q8F6B7	Q8f6b7	leptospira
873	26	65.0	302	10	Q944T9	Q944t9	lycopersico
874	26	65.0	302	16	Q92PB3	Q92pb3	rhizobium m
875	26	65.0	303	16	Q98AQ2	Q98aq2	rhizobium l
876	26	65.0	303	17	Q9HRM8	Q9hrm8	halobacteri
877	26	65.0	305	5	P90872	P90872	caenorhabdi
878	26	65.0	305	11	O70184	O70184	cavia porce
879	26	65.0	305	12	Q8V4U8	Q8v4u8	monkeypox v
880	26	65.0	305	12	Q89518	Q89518	variola vir
881	26	65.0	305	12	Q8JL87	Q8jl87	ectromelia
882	26	65.0	305	12	Q8QMS4	Q8qms4	cowpox viru
883	26	65.0	305	12	O57233	O57233	vaccinia vi
884	26	65.0	305	12	Q80HU5	Q80hu5	vaccinia vi
885	26	65.0	305	12	Q80DU7	Q80du7	cowpox viru
886	26	65.0	306	16	Q82P21	Q82p21	streptomyce
887	26	65.0	307	11	Q8VBX5	Q8vbx5	mus musculu
888	26	65.0	308	10	Q9FY13	Q9fy13	narcissus p
889	26	65.0	308	10	O65882	O65882	populus tri
890	26	65.0	308	10	O65881	O65881	populus tri
891	26	65.0	308	10	O65904	O65904	populus tri
892	26	65.0	310	8	O63625	O63625	dalbulus eb
893	26	65.0	312	16	Q92P56	Q92p56	rhizobium m
894	26	65.0	313	6	Q7YRG4	Q7yrg4	sus scrofa
895	26	65.0	314	17	Q8PUW1	Q8puw1	methanosarc
896	26	65.0	316	13	Q7ZZM8	Q7zzm8	xenopus lae
897	26	65.0	316	16	Q89BW7	Q89bw7	bradyrhizob
898	26	65.0	317	13	Q7T1E1	Q7tle1	gallus gall
899	26	65.0	318	10	Q8GX90	Q8gx90	arabidopsis
900	26	65.0	319	12	Q65409	Q65409	bean common
901	26	65.0	319	12	Q65408	Q65408	bean common
902	26	65.0	319	12	Q65407	Q65407	bean common
903	26	65.0	319	12	Q65410	Q65410	bean common
904	26	65.0	319	12	Q65411	Q65411	bean common
905	26	65.0	320	10	Q7X8Y4	Q7x8y4	oryza sativ
906	26	65.0	320	16	Q8PGJ7	Q8pgj7	xanthomonas
907	26	65.0	321	16	Q9PL79	Q9pl79	chlamydia m
908	26	65.0	321	17	Q8TK75	Q8tk75	methanosarc
909	26	65.0	322	16	Q7VF75	Q7vf75	helicobacte
910	26	65.0	324	16	Q8NUF4	Q8nuf4	staphylococ
911	26	65.0	325	5	O02304	O02304	caenorhabdi
912	26	65.0	326	17	Q9HMZ1	Q9hmz1	halobacteri

913	26	65.0	328	5	Q9VBY3	Q9vby3 drosophila
914	26	65.0	328	5	Q9Txc7	Q9txc7 drosophila
915	26	65.0	332	5	P91231	P91231 caenorhabdi
916	26	65.0	334	4	Q96FS7	Q96fs7 homo sapien
917	26	65.0	335	4	Q96I06	Q96i06 homo sapien
918	26	65.0	335	6	Q863K8	Q863k8 bos taurus
919	26	65.0	337	16	Q8YBY9	Q8yby9 brucella me
920	26	65.0	337	16	Q98DZ8	Q98dz8 rhizobium l
921	26	65.0	337	16	Q8FWD0	Q8fwd0 brucella su
922	26	65.0	338	2	Q8KQV9	Q8kqv9 vibrio chol
923	26	65.0	338	2	Q8RKY1	Q8rkyl providencia
924	26	65.0	338	10	Q40061	Q40061 hordeum vul
925	26	65.0	338	16	Q97IX7	Q97ix7 clostridium
926	26	65.0	338	16	Q8XKG0	Q8xkg0 clostridium
927	26	65.0	338	16	Q88Z62	Q88z62 lactobacill
928	26	65.0	341	16	Q8RBT5	Q8rbt5 thermoanaer
929	26	65.0	341	16	Q81PR6	Q81pr6 bacillus an
930	26	65.0	342	16	O51104	O51104 borrelia bu
931	26	65.0	343	10	Q40148	Q40148 lycopersico
932	26	65.0	343	13	Q803V8	Q803v8 brachydanio
933	26	65.0	344	16	Q98QE5	Q98qe5 mycoplasma
934	26	65.0	347	11	Q8R531	Q8r531 mus musculu
935	26	65.0	347	11	Q8BXN8	Q8bxn8 mus musculu
936	26	65.0	347	16	Q8DI95	Q8di95 synechococc
937	26	65.0	348	5	O61883	O61883 caenorhabdi
938	26	65.0	348	10	O48631	O48631 prunus arme
939	26	65.0	350	6	Q7YRZ9	Q7yrz9 felis silve
940	26	65.0	350	17	Q9V248	Q9v248 pyrococcus
941	26	65.0	352	16	Q8U845	Q8u845 agrobacteri
942	26	65.0	353	4	Q96IC3	Q96ic3 homo sapien
943	26	65.0	354	12	Q65168	Q65168 african swi
944	26	65.0	356	17	Q9HQA7	Q9hqa7 halobacteri
945	26	65.0	357	16	P74141	P74141 synechocyst
946	26	65.0	360	16	Q92JY5	Q92jy5 rhizobium m
947	26	65.0	361	10	Q8LM30	Q8lm30 oryza sativ
948	26	65.0	363	16	Q93SH6	Q93sh6 bradyrhizob
949	26	65.0	366	16	Q92N14	Q92n14 rhizobium m
950	26	65.0	368	5	Q9Y018	Q9y018 reticulomyx
951	26	65.0	368	16	Q8UJX4	Q8ujx4 agrobacteri
952	26	65.0	370	16	Q8XK23	Q8xk23 clostridium
953	26	65.0	374	3	Q9P3B1	Q9p3b1 schizosacch
954	26	65.0	377	16	Q89GP5	Q89gp5 bradyrhizob
955	26	65.0	378	16	Q8XLC0	Q8xlc0 clostridium
956	26	65.0	379	8	Q9G945	Q9g945 microcebus
957	26	65.0	379	8	Q9G944	Q9g944 microcebus
958	26	65.0	379	8	Q9G946	Q9g946 propithecus
959	26	65.0	379	16	Q8R790	Q8r790 thermoanaer
960	26	65.0	381	5	Q09585	Q09585 caenorhabdi
961	26	65.0	382	5	Q9V7J0	Q9v7j0 drosophila
962	26	65.0	386	16	Q9A2F4	Q9a2f4 caulobacter
963	26	65.0	387	2	Q9KI10	Q9ki10 bacillus ce
964	26	65.0	387	16	Q81SK9	Q81sk9 bacillus an
965	26	65.0	389	16	Q89G97	Q89g97 bradyrhizob
966	26	65.0	390	4	Q9H6K3	Q9h6k3 homo sapien
967	26	65.0	390	4	Q96BR8	Q96br8 homo sapien
968	26	65.0	393	8	P92324	P92324 octomeria g
969	26	65.0	397	16	Q99XC4	Q99xc4 staphylococ

970	26	65.0	398	2	O87160	O87160 vibrio chol
971	26	65.0	398	16	Q9WZL9	Q9wzl9 thermotoga
972	26	65.0	398	16	Q9KY68	Q9ky68 streptomyce
973	26	65.0	398	16	Q82HP4	Q82hp4 streptomyce
974	26	65.0	399	2	Q8GF71	Q8gf71 photorhabdu
975	26	65.0	399	4	Q9BPY2	Q9bpy2 homo sapien
976	26	65.0	400	5	Q95WD8	Q95wd8 phlebotomus
977	26	65.0	400	10	Q9XG75	Q9xg75 nicotiana t
978	26	65.0	404	16	Q9CPI6	Q9cpi6 pasteurella
979	26	65.0	404	16	Q87R71	Q87r71 vibrio para
980	26	65.0	408	13	Q9DFS5	Q9dfs5 gallus gall
981	26	65.0	408	16	Q7WF54	Q7wf54 bordetella
982	26	65.0	408	16	Q7W3S6	Q7w3s6 bordetella
983	26	65.0	408	16	Q7VSW3	Q7vsw3 bordetella
984	26	65.0	410	10	Q9ZSY2	Q9zsy2 arabidopsis
985	26	65.0	410	11	Q8C5I3	Q8c5i3 mus musculu
986	26	65.0	413	10	Q94H79	Q94h79 oryza sativ
987	26	65.0	416	11	Q8BY71	Q8by71 mus musculu
988	26	65.0	419	6	O19134	O19134 oryctolagus
989	26	65.0	419	16	O51406	O51406 borrelia bu
990	26	65.0	420	2	Q52501	Q52501 burkholderi
991	26	65.0	422	2	Q8GJ80	Q8gj80 mycobacteri
992	26	65.0	426	10	Q84QA8	Q84qa8 oryza sativ
993	26	65.0	427	11	Q8C0N4	Q8c0n4 mus musculu
994	26	65.0	429	4	Q8NI33	Q8ni33 homo sapien
995	26	65.0	430	2	Q8KZ44	Q8kz44 uncultured
996	26	65.0	433	12	Q84483	Q84483 paramecium
997	26	65.0	436	2	Q9F7N6	Q9f7n6 gamma-prote
998	26	65.0	438	11	O35934	O35934 mus musculu
999	26	65.0	438	11	Q8C5B1	Q8c5b1 mus musculu
1000	26	65.0	439	4	Q9BUG9	Q9bug9 homo sapien

# ALIGNMENTS

## RESULT 1

Q9UCD1

ID Q9UCD1 PRELIMINARY; PRT; 28 AA.  
AC Q9UCD1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Beta-amyloid peptide (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94045685; PubMed=8229004;  
RA Vigo-Pelfrey C., Lee D., Keim P., Lieberburg I., Schenk D.B.;  
RT "Characterization of beta-amyloid peptide from human cerebrospinal  
RT fluid.";  
RL J. Neurochem. 61:1965-1968(1993).  
DR HSSP; P05067; 1AMB.  
DR GO; GO:0016020; C:membrane; IEA.

DR InterPro; IPR001255; Beta-APP.  
DR Pfam; PF03494; Beta-APP; 1.  
SQ SEQUENCE 28 AA; 3244 MW; DE7BD081160AFC81 CRC64;

Query Match 100.0%; Score 40; DB 4; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | | | | |  
Db 16 KLVFFAED 23

#### RESULT 2

Q9UCA9

ID Q9UCA9 PRELIMINARY; PRT; 30 AA.  
AC Q9UCA9;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Beta-amyloid protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94153015; PubMed=8109908;  
RA Wisniewski T., Lalowski M., Levy E., Marques M.R., Frangione B.;  
RT "The amino acid sequence of neuritic plaque amyloid from a familial  
RT Alzheimer's disease patient."  
RL Ann. Neurol. 35:245-246(1994).  
DR HSSP; P05067; 1BA4.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR001255; Beta-APP.  
DR Pfam; PF03494; Beta-APP; 1.  
SQ SEQUENCE 30 AA; 3391 MW; FF4167ABD081160A CRC64;

Query Match 100.0%; Score 40; DB 4; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | | | | |  
Db 16 KLVFFAED 23

#### RESULT 3

Q9UC33

ID Q9UC33 PRELIMINARY; PRT; 33 AA.  
AC Q9UC33;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Beta-amyloid peptide (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93024877; PubMed=1406936;  
 RA Seubert P., Vigo-Pelfrey C., Esch F., Lee M., Dovey H., Davis D.,  
 RA Sinha S., Schlossmacher M., Whaley J., Swindlehurst C.;  
 RT "Isolation and quantification of soluble Alzheimer's beta-peptide from  
 RT biological fluids.";  
 RL Nature 359:325-327(1992).  
 DR HSSP; P05067; 1BA4.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 SQ SEQUENCE 33 AA; 3674 MW; B1DEF2F4167ABD0 CRC64;

Query Match 100.0%; Score 40; DB 4; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 16 KLVFFAED 23

RESULT 4  
 O35463

ID. O35463 PRELIMINARY; PRT; 79 AA.  
 AC O35463;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Alzheimer's amyloid beta protein (Fragment).  
 GN BETA APP.  
 OS Cricetulus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus.  
 OX NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sambamurti K., Pinnix I., Gandhi S.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF030413; AAB86608.1; -.  
 DR HSSP; P05067; 1BA4.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 79 79  
 SQ SEQUENCE 79 AA; 8538 MW; 37F2C6C3BFF3F597 CRC64;

Query Match 100.0%; Score 40; DB 11; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 0.58;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8



|||||||  
Db 36 KLVFFAED 43

RESULT 5

Q16020

ID Q16020 PRELIMINARY; PRT; 82 AA.  
AC Q16020;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Beta-amyloid peptide (Fragment).  
GN BETA APP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93236601; PubMed=8476439;  
RA Denman R.B., Rosenzwaig R., Miller D.L.;  
RT "A system for studying the effect(s) of familial Alzheimer disease  
RT mutations on the processing of the beta-amyloid peptide precursor.";  
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).  
DR EMBL; S61383; AAB26265.2; -.  
DR HSSP; P05067; 1BA4.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR001255; Beta-APP.  
DR Pfam; PF03494; Beta-APP; 1.  
FT NON\_TER 1 1  
FT NON\_TER 82 82  
SQ SEQUENCE 82 AA; 8882 MW; F534AA5AE5D9230A CRC64;

Query Match 100.0%; Score 40; DB 4; Length 82;  
Best Local Similarity 100.0%; Pred. No. 0.61;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 33 KLVFFAED 40

RESULT 6

Q16014

ID Q16014 PRELIMINARY; PRT; 82 AA.  
AC Q16014;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Beta-amyloid peptide (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93236601; PubMed=8476439;

RA Denman R.B., Rosenzwaig R., Miller D.L.;  
 RT "A system for studying the effect(s) of familial Alzheimer disease  
 RT mutations on the processing of the beta-amyloid peptide precursor."  
 RL Biochem. Biophys. Res. Commun. 192:96-103(1993).  
 DR EMBL; S60721; AAB26263.2; -.  
 DR HSSP; P05067; 1BA4.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 82 82  
 SQ SEQUENCE 82 AA; 8972 MW; F534AA5B3EA9230A CRC64;

Query Match 100.0%; Score 40; DB 4; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 0.61;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||||  
 Db 33 KLVFFAED 40

#### RESULT 7

Q16019

ID Q16019 PRELIMINARY; PRT; 82 AA.  
 AC Q16019;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Beta-amyloid peptide (Fragment).  
 GN BETA APP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93236601; PubMed=8476439;  
 RA Denman R.B., Rosenzwaig R., Miller D.L.;  
 RT "A system for studying the effect(s) of familial Alzheimer disease  
 RT mutations on the processing of the beta-amyloid peptide precursor."  
 RL Biochem. Biophys. Res. Commun. 192:96-103(1993).  
 DR EMBL; S61380; AAB26264.2; -.  
 DR HSSP; P05067; 1BA4.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 82 82  
 SQ SEQUENCE 82 AA; 8938 MW; F534AA50E579230A CRC64;

Query Match 100.0%; Score 40; DB 4; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 0.61;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||||

Db 33 KLVFFAED 40

RESULT 8

Q8JH58

ID Q8JH58 PRELIMINARY; PRT; 113 AA.  
AC Q8JH58;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Amyloid beta protein (Fragment).  
OS Chelydra serpentina serpentina (common snapping turtle).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Testudines; Cryptodira; Testudinoidea; Chelydridae; Chelydra.  
OX NCBI\_TaxID=134619;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21876906; PubMed=11882478;  
RA Trudeau V.L., Chiu S., Kennedy S.W., Brooks R.J.;  
RT "Octylphenol (OP) alters the expression of members of the amyloid  
RT protein family in the hypothalamus of the snapping turtle, Chelydra  
RT serpentina serpentina.";  
RL Environ. Health Perspect. 110:269-275(2002).  
DR EMBL; AF541917; AAN04908.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR008155; A4\_APP.  
DR InterPro; IPR001255; Beta-APP.  
DR Pfam; PF03494; Beta-APP; 1.  
DR PRINTS; PR00203; AMYLOIDA4.  
DR PROSITE; PS00320; A4\_INTRA; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 113 AA; 12750 MW; 72515C930496E053 CRC64;

Query Match 100.0%; Score 40; DB 13; Length 113;  
Best Local Similarity 100.0%; Pred. No. 0.83;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
|||||||  
Db 30 KLVFFAED 37

RESULT 9

Q8BPV5

ID Q8BPV5 PRELIMINARY; PRT; 218 AA.  
AC Q8BPV5;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Amyloid beta (Fragment).  
GN APP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";   
 RL Nature 420:563-573(2002).  
 DR EMBL; AK052448; BAC34997.1; -.  
 DR MGD; MGI:88059; App.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 218 AA; 24118 MW; 95B55AFDAE1D0EF5 CRC64;

Query Match 100.0%; Score 40; DB 11; Length 218;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 135 KLVFFAED 142

# RESULT 10

Q8UUI8

ID Q8UUI8 PRELIMINARY; PRT; 357 AA.  
 AC Q8UUI8;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Putative mebrane protein (Fragment).  
 GN APPA.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX PubMed=11862463;  
 RA Musa A., Lehrach H., Russo V.E.A.;  
 RT "Distinct expression patterns of two zebrafish homologues of the human  
 RT APP gene during embryonic development.";   
 RL Dev. Genes Evol. 211:563-567(2001).  
 DR EMBL; AJ315637; CAC85734.1; -.  
 DR ZFIN; ZDB-GENE-000616-13; appa.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PROSITE; PS00320; A4\_INTRA; 1.

FT NON\_TER 1 1  
SQ SEQUENCE 357 AA; 40962 MW; 07D99EEF6C55B2D8 CRC64;

Query Match 100.0%; Score 40; DB 13; Length 357;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 274 KLVFFAED 281

RESULT 11

Q8BPC7

ID Q8BPC7 PRELIMINARY; PRT; 384 AA.  
AC Q8BPC7;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Amyloid beta (Fragment).  
GN APP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
DR EMBL; AK076506; BAC36369.1; -.  
DR MGD; MGI:88059; App.  
DR GO; GO:0005515; F:protein binding; IPI.  
DR InterPro; IPR008155; A4\_APP.  
DR InterPro; IPR001255; Beta-APP.  
DR Pfam; PF03494; Beta-APP; 1.  
DR PRINTS; PR00203; AMYLOIDA4.  
DR PROSITE; PS00320; A4\_INTRA; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 384 AA; 43990 MW; A81B1AD8AE683173 CRC64;

Query Match 100.0%; Score 40; DB 11; Length 384;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 301 KLVFFAED 308

RESULT 12

Q8UUS0

ID Q8UUS0 PRELIMINARY; PRT; 472 AA.

AC Q8UUS0;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Putative membrane protein (Fragment).  
 GN APPA.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX PubMed=11862463;  
 RA Musa A., Lehrach H., Russo V.E.A.;  
 RT "Distinct expression patterns of two zebrafish homologues of the human  
 RT APP gene during embryonic development.";  
 RL Dev. Genes Evol. 211:563-567(2001).  
 DR EMBL; AJ315636; CAC85733.1; -.  
 DR ZFIN; ZDB-GENE-000616-13; appa.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 FT NON TER 1 1  
 SQ SEQUENCE 472 AA; 53787 MW; 24F7128BE3356550 CRC64;

Query Match 100.0%; Score 40; DB 13; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 389 KLVFFAED 396

# RESULT 13

O93296

ID O93296 PRELIMINARY; PRT; 534 AA.  
 AC O93296;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Amyloid protein (Fragment).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98337885; PubMed=9671674;  
 RA Barnes N.Y., Li L., Yoshikawa K., Schwartz L.M., Oppenheim R.W.,  
 RA Milligan C.E.;

RT "Increased production of amyloid precursor protein provides a  
 RT substrate for caspase-3 in dying motoneurons.";  
 RL J. Neurosci. 18:5869-5880(1998).  
 DR EMBL; AF042098; AAC25052.1; -.  
 DR HSSP; P05067; 1BA4.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 534 AA; 60597 MW; FB53ECC2E66D4C92 CRC64;

Query Match 100.0%; Score 40; DB 13; Length 534;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 451 KLVFFAED 458

#### RESULT 14

##### Q9PVL1

ID Q9PVL1 PRELIMINARY; PRT; 569 AA.  
 AC Q9PVL1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Amyloid protein (Fragment).  
 GN APP.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Coulson E.J., Paliga K., Beyreuther K., Masters C.L.;  
 RT "What the evolution of the amyloid protein precursor supergene family  
 RT tells us about its function.";  
 RL Neurochem. Int. 0:0-0(2000).  
 DR EMBL; AF030341; AAF12698.1; -.  
 DR HSSP; P05067; 1BA4.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.

DR PROSITE; PS00320; A4\_INTRA; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 569 AA; 64753 MW; 0AB8BB851863A19D CRC64;

Query Match 100.0%; Score 40; DB 13; Length 569;  
Best Local Similarity 100.0%; Pred. No. 4.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
|||||||  
Db 487 KLVFFAED 494

#### RESULT 15

Q9I9E7

ID Q9I9E7 PRELIMINARY; PRT; 612 AA.  
AC Q9I9E7;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Amyloid protein (Fragment).  
GN APPA.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Slavov D.B., Gardiner K.;  
RT "An App cDNA from Zebrafish (Danio rerio).";  
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF257742; AAF71748.1; -.  
DR HSSP; P05067; 1HZ3.  
DR ZFIN; ZDB-GENE-000616-13; appa.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR008155; A4\_APP.  
DR InterPro; IPR008154; A4\_extra.  
DR InterPro; IPR001255; Beta-APP.  
DR Pfam; PF02177; A4\_EXTRA; 1.  
DR Pfam; PF03494; Beta-APP; 1.  
DR PRINTS; PR00203; AMYLOIDA4.  
DR PROSITE; PS00319; A4\_EXTRA; 1.  
DR PROSITE; PS00320; A4\_INTRA; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 612 AA; 69710 MW; 59A9ACBDF9C59EFF CRC64;

Query Match 100.0%; Score 40; DB 13; Length 612;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
|||||||  
Db 529 KLVFFAED 536

#### RESULT 16



Q7ZZT1

ID Q7ZZT1 PRELIMINARY; PRT; 678 AA.  
AC Q7ZZT1;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Amyloid protein a variant 2.  
GN APPA.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Groth C., Lardelli M.;  
RT "Investigation of zebrafish appa expression during embryogenesis."  
RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AY271746; AAP22958.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR008155; A4\_APP.  
DR InterPro; IPR008154; A4\_extra.  
DR InterPro; IPR001255; Beta-APP.  
DR Pfam; PF02177; A4\_EXTRA; 1.  
DR Pfam; PF03494; Beta-APP; 1.  
DR PRINTS; PR00203; AMYLOIDA4.  
DR SMART; SM00006; A4\_EXTRA; 1.  
DR PROSITE; PS00319; A4\_EXTRA; 1.  
DR PROSITE; PS00320; A4\_INTRA; 1.  
SQ SEQUENCE 678 AA; 76755 MW; 94163778444FD0BC CRC64;

Query Match 100.0%; Score 40; DB 13; Length 678;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | | | | |  
Db 595 KLVFFAED 602

RESULT 17

Q9DGJ8

ID Q9DGJ8 PRELIMINARY; PRT; 695 AA.  
AC Q9DGJ8;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Beta-amyloid precursor protein 695 isoform.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sarasa M., Rodolosse A., Sorribas V.;  
RT "Cloning of full-length chicken beta-amyloid precursor protein

RT isoforms.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF289218; AAG00593.1; -.  
 DR HSSP; P05067; 1BA4.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 SQ SEQUENCE 695 AA; 78565 MW; F201ED02AEC86D95 CRC64;

Query Match 100.0%; Score 40; DB 13; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 612 KLVFFAED 619

RESULT 18  
 Q90W28

ID Q90W28 PRELIMINARY; PRT; 738 AA.  
 AC Q90W28;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Amyloid precursor protein.  
 GN APPA OR APP.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Groth C., Lardelli M.;  
 RT "Expression analysis of zebrafish app.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF389401; AAK64495.1; -.  
 DR ZFIN; ZDB-GENE-000616-13; appa.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR Pfam; PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PRINTS; PR00759; BASICPTASE.

DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 KW Protease inhibitor; Serine protease inhibitor.  
 SQ SEQUENCE 738 AA; 83577 MW; AF480F6D308FD298 CRC64;

Query Match 100.0%; Score 40; DB 13; Length 738;  
 Best Local Similarity 100.0%; Pred. No. 5.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||  
 Db 655 KLVFFAED 662

# RESULT 19

Q9DGJ7

ID Q9DGJ7 PRELIMINARY; PRT; 751 AA.  
 AC Q9DGJ7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Beta-amyloid precursor protein 751 isoform.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sarasa M., Rodolosse A., Sorribas V.;  
 RT "Cloning of full-length chicken beta-amyloid precursor protein  
 RT isoforms."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF289219; AAG00594.1; -.  
 DR HSSP; P05067; 1BA4.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR Pfam; PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.

DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
KW Protease inhibitor; Serine protease inhibitor.  
SQ SEQUENCE 751 AA; 84705 MW; E78E9413A8033D84 CRC64;

Query Match 100.0%; Score 40; DB 13; Length 751;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|||||||  
Db 668 KLVFFAED 675

RESULT 20

Q8A6R7

ID Q8A6R7 PRELIMINARY; PRT; 1676 AA.  
AC Q8A6R7;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Conserved protein, with weak BamHI domain.  
GN BT1809.  
OS Bacteroides thetaiotaomicron.  
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
OC Bacteroidaceae; Bacteroides.  
OX NCBI\_TaxID=818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VPI-5482 / ATCC 29148;  
RX MEDLINE=22550858; PubMed=12663928;  
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
RA Chiang H.C., Hooper L.V., Gordon J.I.;  
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."  
RL Science 299:2074-2076(2003).  
DR EMBL; AE016933; AAO76916.1; -.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR002048; EF-hand.  
DR PROSITE; PS00018; EF\_HAND; 1.  
KW Complete proteome.  
SQ SEQUENCE 1676 AA; 193674 MW; 28065878C0F6C961 CRC64;

Query Match 92.5%; Score 37; DB 16; Length 1676;  
Best Local Similarity 75.0%; Pred. No. 52;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::|||||  
Db 1656 KIIFFAED 1663

RESULT 21

Q98SG0

ID Q98SG0 PRELIMINARY; PRT; 693 AA.  
AC Q98SG0;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Beta-amyloid precursor protein A.  
 GN APP.  
 OS *Xenopus laevis* (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; *Xenopus*.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Van den Hurk W.H.;  
 RL Thesis (2001), Department of Biological Sciences,  
 RL University of Nijmegen, Nijmegen, Netherlands.  
 DR EMBL; AJ298150; CAC37193.1; -.  
 DR HSSP; P05067; 1HZ3.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 KW Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 SQ SEQUENCE 693 AA; 78568 MW; CAF1DF655C1AB653 CRC64;

Query Match 90.0%; Score 36; DB 13; Length 693;  
 Best Local Similarity 87.5%; Pred. No. 36;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||||:  
 Db 610 KLVFFAEE 617

# RESULT 22

Q91963

ID Q91963 PRELIMINARY; PRT; 747 AA.  
 AC Q91963;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE APP747.  
 GN APP747.  
 OS *Xenopus*.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae.  
 OX NCBI\_TaxID=8353;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93129227; PubMed=1282805;  
 RA Okado H., Okamoto H.;  
 RT "A *Xenopus* homologue of the human beta-amyloid precursor protein:

RT developmental regulation of its gene expression.";  
 RL Biochem. Biophys. Res. Commun. 189:1561-1568(1992).  
 DR EMBL; S52417; AAB24853.1; -.  
 DR HSSP; P05067; 1HZ3.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR InterPro; IPR002223; Kunitz\_BPTI.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR Pfam; PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Kunitz\_BPTI; 1.  
 DR SMART; SM00131; KU; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 1.  
 KW Protease inhibitor; Serine protease inhibitor.  
 SQ SEQUENCE 747 AA; 84893 MW; A75E81885681D948 CRC64;

Query Match 90.0%; Score 36; DB 13; Length 747;  
 Best Local Similarity 87.5%; Pred. No. 38;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||:  
 Db 664 KLVFFAEE 671

# RESULT 23

Q32406

ID Q32406 PRELIMINARY; PRT; 162 AA.  
 AC Q32406;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Heteranthera limosa.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Heteranthera.  
 OX NCBI\_TaxID=44958;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae.";  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41608; AAD09846.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.

DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18679 MW; 1EDB2B1BB8ADB548 CRC64;

Query Match 82.5%; Score 33; DB 8; Length 162;  
 Best Local Similarity 62.5%; Pred. No. 37;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |::||:|  
 Db 13 KVIFFSED 20

#### RESULT 24

Q7VR77

ID Q7VR77 PRELIMINARY; PRT; 197 AA.  
 AC Q7VR77;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Peptidyl-tRNA hydrolase (EC 3.1.1.29).  
 GN PTH OR BFL345.  
 OS Candidatus Blochmannia floridanus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.  
 OX NCBI\_TaxID=203907;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22784745; PubMed=12886019;  
 RA Gil R., Silva F.J., Zientz E., Delmotte F., Gonzalez-Candelas F.,  
 RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,  
 RA van Ham R.C.H.J., Gross R., Moya A.;  
 RT "The genome sequence of Blochmannia floridanus: comparative analysis  
 RT of reduced genomes."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393(2003).  
 DR EMBL; BX248585; CAD83412.1; -.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 197 AA; 22610 MW; 251584AD1EF7BE6E CRC64;

Query Match 82.5%; Score 33; DB 16; Length 197;  
 Best Local Similarity 75.0%; Pred. No. 45;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |:||||  
 Db 169 KVVFFTED 176

#### RESULT 25

Q98SF9

ID Q98SF9 PRELIMINARY; PRT; 695 AA.  
 AC Q98SF9;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Beta-amyloid precursor protein B.  
 GN APP.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipioidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Van den Hurk W.H.;  
 RL Thesis (2001), Department of Biological Sciences,  
 RL University of Nijmegen, Nijmegen, Netherlands.  
 DR EMBL; AJ298151; CAC37194.1; -.  
 DR HSSP; P05067; 1HZ3.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 KW Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 SQ SEQUENCE 695 AA; 78803 MW; DC14EB02AFB0204A CRC64;  
  
 Query Match 82.5%; Score 33; DB 13; Length 695;  
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 KLVFFAED 8  
 |||||:::  
 Db 612 KLVFFADE 619

# RESULT 26

Q7ZXQ0

ID Q7ZXQ0 PRELIMINARY; PRT; 695 AA.  
 AC Q7ZXQ0;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipioidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.;



RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC044324; AAH44324.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.  
 DR SMART; SM00006; A4\_EXTRA; 1.  
 DR PROSITE; PS00319; A4\_EXTRA; 1.  
 DR PROSITE; PS00320; A4\_INTRA; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 695 AA; 78803 MW; C1BD8AACC3356B05 CRC64;

Query Match 82.5%; Score 33; DB 13; Length 695;  
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||||::  
 Db 612 KLVFFADE 619

# RESULT 27

Q9Z588

ID Q9Z588 PRELIMINARY; PRT; 182 AA.  
 AC Q9Z588;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE SC8D9.03 protein.  
 GN SCO5491 OR SC8D9.03.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Murphy L., Harris D.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]

RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,  
 RA Rabbinoiwitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete *Streptomyces*  
 RT *coelicolor* A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL; AL939123; CAB37567.1; -.  
 DR PIR; T35807; T35807.  
 DR InterPro; IPR005269; Cons\_hypoth730.  
 DR Pfam; PF03641; Lysine\_decarbox; 1.  
 DR TIGRFAMs; TIGR00730; TIGR00730; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 182 AA; 19721 MW; 8494A6107A939E49 CRC64;

Query Match 80.0%; Score 32; DB 16; Length 182;  
 Best Local Similarity 75.0%; Pred. No. 69;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 :|||||:  
 Db 160 ELVFFAEE 167

# RESULT 28

Q931V3

ID Q931V3 PRELIMINARY; PRT; 184 AA.  
 AC Q931V3;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE GTP pyrophosphokinase.  
 GN YJBM OR SAV1001.  
 OS *Staphylococcus aureus* (strain Mu50 / ATCC 700699).  
 OC Bacteria; Firmicutes; Bacillales; *Staphylococcus*.  
 OX NCBI\_TaxID=158878;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21311952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
 RT "Whole genome sequencing of *meticillin*-resistant *Staphylococcus*  
 RT *aureus*.";  
 RL Lancet 357:1225-1240(2001).

DR EMBL; AP003361; BAB57163.1; -.  
DR GO; GO:0016301; F:kinase activity; IEA.  
KW Kinase; Complete proteome.  
SQ SEQUENCE 184 AA; 21755 MW; 347AD094AFCCCD03 CRC64;

Query Match 80.0%; Score 32; DB 16; Length 184;  
Best Local Similarity 85.7%; Pred. No. 69;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
|||:|  
Db 81 LVFFSED 87

RESULT 29

Q7X225

ID Q7X225 PRELIMINARY; PRT; 261 AA.  
AC Q7X225;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE GTP pyrophosphokinase.  
GN YIBM.  
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=158878;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wootton M., Avison M.B., Bennett P.M., Howe R.A., MacGowan A.P.,  
RA Walsh T.R.;  
RT "Genetic analysis of seventeen genes in Staphylococcus aureus with  
RT reduced susceptibility to vancomycin (VRSA) and hetero-VRSA (hVRSA).";  
RL Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ564070; CAD91854.1; -.  
KW Kinase.  
SQ SEQUENCE 261 AA; 30592 MW; CE58B6DE0F263166 CRC64;

Query Match 80.0%; Score 32; DB 2; Length 261;  
Best Local Similarity 85.7%; Pred. No. 98;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
|||:|  
Db 158 LVFFSED 164

RESULT 30

Q7WRM0

ID Q7WRM0 PRELIMINARY; PRT; 261 AA.  
AC Q7WRM0;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE GTP pyrophosphokinase.  
GN YIBM.  
OS Staphylococcus aureus.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mu3, Michigan, New Jersey, Glasgow3700, Glasgow3759, PC3, St  
 RC Luke, Sweden307, LIM2, LIM3, PC1, Sweden309, FranceDuf, Slovenia6096,  
 RC Southampton23, LIM1, Norway1018, SMH2, LiverpoolAG, LLA, LLE,  
 RC SMH11888, SMH18000, SMH18034, SMH18037, SMH12248, SMH8997, SMH10501,  
 RC SMH17608, SMH17487, and SMH14017;  
 RA Wootton M., Avison M.B., Bennett P.M., Howe R.A., MacGowan A.P.,  
 RA Walsh T.R.;  
 RT "Genetic analysis of seventeen genes in Staphylococcus aureus with  
 RT reduced susceptibility to vancomycin (VRSA) and hetero-VRSA (hVRSA).";  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AJ564071; CAD91855.1; -.  
 DR EMBL; AJ564072; CAD91856.1; -.  
 DR EMBL; AJ564073; CAD91857.1; -.  
 DR EMBL; AJ564074; CAD91858.1; -.  
 DR EMBL; AJ564075; CAD91859.1; -.  
 DR EMBL; AJ564076; CAD91860.1; -.  
 DR EMBL; AJ564077; CAD91861.1; -.  
 DR EMBL; AJ564078; CAD91862.1; -.  
 DR EMBL; AJ564079; CAD91863.1; -.  
 DR EMBL; AJ564080; CAD91864.1; -.  
 DR EMBL; AJ564081; CAD91865.1; -.  
 DR EMBL; AJ564082; CAD91866.1; -.  
 DR EMBL; AJ564083; CAD91867.1; -.  
 DR EMBL; AJ564084; CAD91868.1; -.  
 DR EMBL; AJ564085; CAD91869.1; -.  
 DR EMBL; AJ564086; CAD91870.1; -.  
 DR EMBL; AJ564087; CAD91871.1; -.  
 DR EMBL; AJ564088; CAD91872.1; -.  
 DR EMBL; AJ564089; CAD91873.1; -.  
 DR EMBL; AJ564090; CAD91874.1; -.  
 DR EMBL; AJ564091; CAD91875.1; -.  
 DR EMBL; AJ564092; CAD91876.1; -.  
 DR EMBL; AJ564093; CAD91877.1; -.  
 DR EMBL; AJ564094; CAD91878.1; -.  
 DR EMBL; AJ564095; CAD91879.1; -.  
 DR EMBL; AJ564096; CAD91880.1; -.  
 DR EMBL; AJ564097; CAD91881.1; -.  
 DR EMBL; AJ564098; CAD91882.1; -.  
 DR EMBL; AJ564099; CAD91883.1; -.  
 DR EMBL; AJ564100; CAD91884.1; -.  
 DR EMBL; AJ564101; CAD91885.1; -.  
 KW Kinase.  
 SQ SEQUENCE 261 AA; 30592 MW; CE58B6DE0F263166 CRC64;  
  
 Query Match 80.0%; Score 32; DB 2; Length 261;  
 Best Local Similarity 85.7%; Pred. No. 98;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 2 LVFFAED 8  
 ||||:||  
 Db 158 LVFFSED 164

Q99V89

ID Q99V89 PRELIMINARY; PRT; 261 AA.  
 AC Q99V89;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein SA0860.  
 GN SA0860.  
 OS Staphylococcus aureus (strain N315).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=158879;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21311952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,  
 RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,  
 RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,  
 RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,  
 RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,  
 RA Ogasawara N., Hayashi H., Hiramatsu K.;  
 RT "Whole genome sequencing of meticillin-resistant Staphylococcus  
 RT aureus.";  
 RL Lancet 357:1225-1240(2001).  
 DR EMBL; AP003132; BAB42101.1; -.  
 DR PIR; B89868; B89868.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 261 AA; 30592 MW; CE58B6DE0F263166 CRC64;

Query Match 80.0%; Score 32; DB 16; Length 261;  
 Best Local Similarity 85.7%; Pred. No. 98;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
 ||||:||  
 Db 158 LVFFSED 164

RESULT 32

Q8NXD0

ID Q8NXD0 PRELIMINARY; PRT; 268 AA.  
 AC Q8NXD0;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein MW0883.  
 GN MW0883.  
 OS Staphylococcus aureus (strain MW2).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=196620;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22040717; PubMed=12044378;  
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
 RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,  
 RA Yamamoto K., Hiramatsu K.;  
 RT "Genome and virulence determinants of high virulence community-

RT acquired MRSA.";  
RL Lancet 359:1819-1827(2002).  
DR EMBL; AP004825; BAB94748.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 268 AA; 31363 MW; F40B877860BD1140 CRC64;

Query Match 80.0%; Score 32; DB 16; Length 268;  
Best Local Similarity 85.7%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
||||:|  
Db 165 LVFFSED 171

RESULT 33

Q8CUH9

ID Q8CUH9 PRELIMINARY; PRT; 282 AA.  
AC Q8CUH9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Spore coat polysaccharide synthesis (dTDP-4-dehydrorhamnose reductase).  
GN OB1128.  
OS Oceanobacillus iheyensis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
OX NCBI\_TaxID=182710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;  
RX MEDLINE=22220767; PubMed=12235376;  
RA Takami H., Takaki Y., Uchiyama I.;  
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya  
RT Ridge and its unexpected adaptive capabilities to extreme  
RT environments.";  
RL Nucleic Acids Res. 30:3927-3935(2002).  
DR EMBL; AP004596; BAC13084.1; -.  
DR GO; GO:0008831; F:dTDP-4-dehydrorhamnose reductase activity; IEA.  
DR GO; GO:0045226; P:extracellular polysaccharide biosynthesis; IEA.  
DR InterPro; IPR005913; TDP\_rham\_reduct.  
DR Pfam; PF04321; RmlD\_sub\_bind; 1.  
DR TIGRFAMs; TIGR01214; rmlD; 1.  
KW Complete proteome.  
SQ SEQUENCE 282 AA; 32214 MW; F53A8BE4CCD17075 CRC64;

Query Match 80.0%; Score 32; DB 16; Length 282;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||||:|  
Db 98 KLVFFSSD 105

RESULT 34

Q10833

ID Q10833 PRELIMINARY; PRT; 472 AA.  
 AC Q10833;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Collagenase-3.  
 GN GENE A.  
 OS *Xenopus laevis* (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; *Xenopus*.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96312514; PubMed=8700860;  
 RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,  
 RA Remo B.F., Pinder A.;  
 RT "The thyroid hormone-induced tail resorption program during *Xenopus*  
 RT *laevis* metamorphosis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929(1996).  
 DR EMBL; U41824; AAC59870.1; -.  
 DR HSSP; P45452; lPEX.  
 DR MEROPS; M10.013; -.  
 DR GO; GO:0005578; C:extracellular matrix; IEA.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000585; Hemopexin.  
 DR InterPro; IPR006026; Peptidase\_M.  
 DR InterPro; IPR001818; Pept\_M10A\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR Pfam; PF00045; hemopexin; 4.  
 DR Pfam; PF00413; Peptidase\_M10; 1.  
 DR Pfam; PF03933; Peptidase\_M10\_N; 1.  
 DR PRINTS; PR00138; MATRIXIN.  
 DR SMART; SM00120; HX; 4.  
 DR SMART; SM00235; ZnMc; 1.  
 DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
 DR PROSITE; PS00024; HEMOPEXIN; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Hydrolase; Metal-binding; Metalloprotease; Protease; Zinc.  
 SQ SEQUENCE 472 AA; 53837 MW; CDF490F5738593B4 CRC64;  
  
 Query Match 80.0%; Score 32; DB 13; Length 472;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 Qy 1 KLVFFAED 8  
 | :||||  
 Db 394 KTLFFAED 401

RESULT 35

Q7UZZ9

ID Q7UZZ9 PRELIMINARY; PRT; 501 AA.  
 AC Q7UZZ9;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein precursor.  
 GN PMM1566.  
 OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).  
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
 OC Prochlorococcus.  
 OX NCBI\_TaxID=59919;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22825698; PubMed=12917642;  
 RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,  
 RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,  
 RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,  
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
 RA Webb E.A., Zinser E.R., Chisholm S.W.;  
 RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic  
 RT niche differentiation."  
 RL Nature 424:1042-1047(2003).  
 DR EMBL; BX572094; CAE20025.1; -.  
 KW Hypothetical protein; Signal; Complete proteome.  
 FT SIGNAL 1 20 Potential.  
 SQ SEQUENCE 501 AA; 56225 MW; 2300E99F811F0C81 CRC64;

Query Match 80.0%; Score 32; DB 16; Length 501;  
 Best Local Similarity 62.5%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |:|:|:|  
 Db 279 KVMYFADD 286

# RESULT 36

Q9VX31

ID Q9VX31 PRELIMINARY; PRT; 1105 AA.  
 AC Q9VX31;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE CG8211 protein.  
 GN CG8211.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,



RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).

RN [2]

RP SEQUENCE FROM N.A.

RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,

RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AE003506; AAF48750.2; -.  
 DR FlyBase; FBgn0030858; CG8211.  
 SQ SEQUENCE 1105 AA; 124889 MW; 6FC8C810A52D9CA2 CRC64;

Query Match 80.0%; Score 32; DB 5; Length 1105;  
 Best Local Similarity 75.0%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 || ||||:  
 Db 583 KLTFFAAEE 590

#### RESULT 37

##### Q9GYD1

ID Q9GYD1 PRELIMINARY; PRT; 2613 AA.  
 AC Q9GYD1;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Possible dynein heavy chain alpha (Fragment).  
 GN L3557.01.  
 OS Leishmania major.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Friedlin;  
 RA Zimmermann W., Wambutt R., Ivens A.C., Quail M., Rajandream M.A.,  
 RA Barrell B.G.;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Friedlin;  
 RX MEDLINE=98146435; PubMed=9477341;  
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 RA Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome."  
 RL Genome Res. 8:135-145(1998).  
 DR EMBL; AL391562; CAC04268.1; -.  
 DR GO; GO:0030286; C:dynein complex; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.  
 DR GO; GO:0008567; F:dynein ATPase activity; IEA.  
 DR GO; GO:0003777; F:microtubule motor activity; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.

DR GO; GO:0007018; P:microtubule-based movement; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR004273; Dynein\_heavy.  
 DR InterPro; IPR000169; SHprot\_acsite.  
 DR Pfam; PF03028; Dynein\_heavy; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
 KW ATP-binding.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 2613 AA; 295773 MW; 0FFD328B61CCE6B0 CRC64;

Query Match 80.0%; Score 32; DB 5; Length 2613;  
 Best Local Similarity 75.0%; Pred. No. 9.5e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||||| :|  
 Db 560 KLVFFLDD 567

RESULT 38  
 O97917

ID O97917 PRELIMINARY; PRT; 49 AA.  
 AC O97917;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Amyloid protein (Fragment).  
 GN APP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20063685; PubMed=10594237;  
 RA Konfortov B.A., Licence V.E., Miller J.R.;  
 RT "Resequencing of DNA from a diverse panel of cattle reveals a high  
 RT level of polymorphism in both intron and exon."  
 RL Mamm. Genome 10:1142-1145(1999).  
 DR EMBL; AJ133033; CAB38017.1; -.  
 DR HSSP; P05067; 1BA4.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF03494; Beta-APP; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 49 49  
 SQ SEQUENCE 49 AA; 5183 MW; 6287463F0559BDED CRC64;

Query Match 77.5%; Score 31; DB 6; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
 |||||

Db 1 VFFAED 6

RESULT 39

Q8A5K5

ID Q8A5K5 PRELIMINARY; PRT; 147 AA.  
AC Q8A5K5;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative cytosine/adenosine deaminase.  
GN BT2234.  
OS Bacteroides thetaiotaomicron.  
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
OC Bacteroidaceae; Bacteroides.  
OX NCBI\_TaxID=818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VPI-5482 / ATCC 29148;  
RX MEDLINE=22550858; PubMed=12663928;  
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
RA Chiang H.C., Hooper L.V., Gordon J.I.;  
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";  
RL Science 299:2074-2076(2003).  
DR EMBL; AE016935; AAO77341.1; -.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.  
DR InterPro; IPR002125; dCMP/cyt\_deam.  
DR Pfam; PF00383; dCMP\_cyt\_deam; 1.  
DR PROSITE; PS00903; CYT\_DCMP\_DEAMINASES; 1.  
KW Complete proteome.  
SQ SEQUENCE 147 AA; 15970 MW; 3A072D8ED1686443 CRC64;

Query Match 77.5%; Score 31; DB 16; Length 147;  
Best Local Similarity 87.5%; Pred. No. 91;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | |  
Db 99 KLVFGAED 106

RESULT 40

Q82JK4

ID Q82JK4 PRELIMINARY; PRT; 179 AA.  
AC Q82JK4;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN SAV2751.  
OS Streptomyces avermitilis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=33903;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=21477403; PubMed=11572948;  
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,  
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
RT "Genome sequence of an industrial microorganism *Streptomyces*  
RT *avermitilis*: deducing the ability of producing secondary  
RT metabolites.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=22608306; PubMed=12692562;  
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
RA Sakaki Y., Hattori M., Omura S.;  
RT "Complete genome sequence and comparative analysis of the industrial  
RT microorganism *Streptomyces avermitilis*.";  
RL Nat. Biotechnol. 21:526-531(2003).  
DR EMBL; AP005032; BAC70462.1; -.  
DR InterPro; IPR005269; Cons\_hypoth730.  
DR Pfam; PF03641; Lysine\_decarbox; 1.  
DR TIGRFAMs; TIGR00730; TIGR00730; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 179 AA; 19286 MW; 75B53B6069659145 CRC64;

Query Match 77.5%; Score 31; DB 16; Length 179;  
Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LVFFAED 8  
|||||:  
Db 158 LVFFAEE 164

#### RESULT 41

##### Q8ESR7

ID Q8ESR7 PRELIMINARY; PRT; 208 AA.  
AC Q8ESR7;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ATP-phosphoribosyltransferase.  
GN OB0552.  
OS *Oceanobacillus iheyensis*.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Oceanobacillus*.  
OX NCBI\_TaxID=182710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;  
RX MEDLINE=22220767; PubMed=12235376;  
RA Takami H., Takaki Y., Uchiyama I.;  
RT "Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya  
RT Ridge and its unexpected adaptive capabilities to extreme  
RT environments.";  
RL Nucleic Acids Res. 30:3927-3935(2002).  
DR EMBL; AP004594; BAC12508.1; -.  
DR GO; GO:0003879; F:ATP phosphoribosyltransferase activity; IEA.

DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.  
 DR GO; GO:0000105; P:histidine biosynthesis; IEA.  
 DR InterPro; IPR001348; ATP\_phospho\_trans.  
 DR Pfam; PF01634; HisG; 1.  
 DR ProDom; PD003516; ATP\_phospho\_trans; 1.  
 DR TIGRFAMs; TIGR00070; hisG; 1.  
 DR PROSITE; PS01316; ATP\_P\_PHORIBOSYLTR; 1.  
 KW Glycosyltransferase; Transferase; Complete proteome.  
 SQ SEQUENCE 208 AA; 22633 MW; 0A32E0CA7D029343 CRC64;

Query Match 77.5%; Score 31; DB 16; Length 208;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 ||||: ||  
 Db 37 KLVFYND 44

# RESULT 42

Q8E2V5

ID Q8E2V5 PRELIMINARY; PRT; 228 AA.  
 AC Q8E2V5;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Gbs2009 protein.  
 GN GBS2009.  
 OS Streptococcus agalactiae (serotype III).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=216495;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEM316 / Serotype III;  
 RX MEDLINE=22242508; PubMed=12354221;  
 RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,  
 RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,  
 RA Kunst F.;  
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing  
 RT invasive neonatal disease."  
 RL Mol. Microbiol. 45:1499-1513(2002).  
 DR EMBL; AL766855; CAD47668.1; -.  
 DR SagaList; gbs2009; -.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0000156; F:two-component response regulator activity; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.  
 DR InterPro; IPR001789; Response\_reg.  
 DR InterPro; IPR001867; Trans\_reg\_C.  
 DR Pfam; PF00072; response\_reg; 1.  
 DR Pfam; PF00486; trans\_reg\_C; 1.  
 DR ProDom; PD000039; Response\_reg; 1.  
 DR ProDom; PD000329; Trans\_reg\_C; 1.  
 DR SMART; SM00448; REC; 1.  
 DR PROSITE; PS50110; RESPONSE\_REGULATORY; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 228 AA; 26617 MW; 5FDE87004BBC6518 CRC64;

Query Match 77.5%; Score 31; DB 16; Length 228;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|:::||:|  
Db 3 KIIYFADD 10

RESULT 43

Q8DX05

ID Q8DX05 PRELIMINARY; PRT; 228 AA.  
AC Q8DX05;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE DNA-binding response regulator.  
GN SAG2054.  
OS Streptococcus agalactiae (serotype V).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=216466;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2603 V/R / Serotype V;  
RX MEDLINE=22222988; PubMed=12200547;  
RA Tettelin H., Massignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,  
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,  
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,  
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,  
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,  
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,  
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,  
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,  
RA Fraser C.M.;  
RT "Complete genome sequence and comparative genomic analysis of an  
RT emerging human pathogen, serotype V Streptococcus agalactiae."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).  
DR EMBL; AE014284; AAN00913.1; -.  
DR TIGR; SAG2054; -.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0000156; F:two-component response regulator activity; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.  
DR InterPro; IPR001789; Response\_reg.  
DR InterPro; IPR001867; Trans\_reg\_C.  
DR Pfam; PF00072; response\_reg; 1.  
DR Pfam; PF00486; trans\_reg\_C; 1.  
DR ProDom; PD000039; Response\_reg; 1.  
DR ProDom; PD000329; Trans\_reg\_C; 1.  
DR SMART; SM00448; REC; 1.  
DR PROSITE; PS50110; RESPONSE\_REGULATORY; 1.  
KW DNA-binding; Complete proteome.  
SQ SEQUENCE 228 AA; 26617 MW; 5FDE87004BBC6518 CRC64;

Query Match 77.5%; Score 31; DB 16; Length 228;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::|||:

Db 3 KIIYFADD 10

RESULT 44

Q8I3W8

ID Q8I3W8 PRELIMINARY; PRT; 248 AA.  
AC Q8I3W8;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN PFE0695W.  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,  
RA Hall N., Bowman S., Churcher C., Quail M., Barrell B.;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22255708; PubMed=12368867;  
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,  
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,  
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,  
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,  
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,  
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,  
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,  
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,  
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,  
RA Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,  
RA Oliver K., Ormond D., Price C., Quail M.A., Rabbinowitsch E.,  
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,  
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,  
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,  
RA Sulston J.E., Craig A., Newbold C., Barrell B.G;  
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";  
RL Nature 419:527-531(2002).  
DR EMBL; AL929352; CAD51504.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 248 AA; 29745 MW; 86FC92C34D631CAD CRC64;

Query Match 77.5%; Score 31; DB 5; Length 248;  
Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::|||:

Db 208 KILFFAKD 215



RESULT 45

Q8NN32

ID Q8NN32 PRELIMINARY; PRT; 259 AA.  
 AC Q8NN32;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Transcriptional regulator.  
 GN CGL2381.  
 OS Corynebacterium glutamicum (Brevibacterium flavum).  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=1718;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
 RA Nakagawa S.;  
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AP005281; BAB99774.1; -.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001647; HTH\_TetR.  
 DR Pfam; PF00440; tetR; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 259 AA; 28363 MW; D487361648C34A2C CRC64;

Query Match 77.5%; Score 31; DB 16; Length 259;  
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
 | |||||  
 Db 37 LTFFAED 43

RESULT 46

Q9IK76

ID Q9IK76 PRELIMINARY; PRT; 290 AA.  
 AC Q9IK76;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE SsDNA binding protein.  
 GN DBP.  
 OS Spodoptera litura nucleopolyhedrovirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 OC Nucleopolyhedrovirus.  
 OX NCBI\_TaxID=46242;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Li C., Pang Y., Yan Q.;  
 RT "Cloning and sequence analysis of gp37, 39k, vubi, dbp gene and other  
 RT two ORFs in the EcoRI-D fragment of Spodoptera litura  
 RT nucleopolyhedrovirus genome.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=G2;  
 RX MEDLINE=21425398; PubMed=11531416;  
 RA Pang Y., Yu J., Wang L., Hu X., Bao W., Li G., Chen C., Han H., Hu S.,  
 RA Yang H.;  
 RT "Sequence Analysis of the Spodoptera litura Multicapsid  
 RT Nucleopolyhedrovirus Genome.";  
 RL Virology 287:391-404(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=G2;  
 RA Yu J., Wang L., Hu X., Pang Y.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF216301; AAF72589.1; -.  
 DR EMBL; AF325155; AAL01716.1; -.  
 DR InterPro; IPR006871; Baculo\_DNA\_bind.  
 DR Pfam; PF04786; Baculo\_DNA\_bind; 1.  
 SQ SEQUENCE 290 AA; 33297 MW; 0EF551C6243A7FFC CRC64;

Query Match 77.5%; Score 31; DB 12; Length 290;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||:| ||  
 Db 31 KLVYFDED 38

#### RESULT 47

Q814E2

ID Q814E2 PRELIMINARY; PRT; 355 AA.  
 AC Q814E2;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical cytosolic protein.  
 GN BCP0010.  
 OS Bacillus cereus (strain ATCC 14579 / DSM 31).  
 OG Plasmid pBClin15.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=226900;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22608415; PubMed=12721630;  
 RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,  
 RA Kapatral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,  
 RA Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,  
 RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,  
 RA Overbeek R., Kyrpides N.;  
 RT "Genome sequence of Bacillus cereus and comparative analysis with  
 RT Bacillus anthracis.";  
 RL Nature 423:87-91(2003).  
 DR EMBL; AE016878; AAP12354.1; -.  
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 KW Hypothetical protein; Plasmid; Complete proteome.  
 SQ SEQUENCE 355 AA; 39842 MW; 9D1AA76AA40813AF CRC64;

Query Match 77.5%; Score 31; DB 16; Length 355;  
Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||:|: ||  
Db 248 KLIFYIED 255

RESULT 48

Q7WD43

ID Q7WD43 PRELIMINARY; PRT; 370 AA.  
AC Q7WD43;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative integral membrane protein.  
GN BB3735.  
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Alcaligenaceae; Bordetella.  
OX NCBI\_TaxID=518;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RB50 / ATCC BAA-588;  
RX MEDLINE=22827954; PubMed=12910271;  
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,  
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,  
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
RT Bordetella parapertussis and Bordetella bronchiseptica."  
RL Nat. Genet. 35:32-40(2003).  
DR EMBL; BX640448; CAE35709.1; -.  
KW Complete proteome.  
SQ SEQUENCE 370 AA; 40050 MW; C9E4D8123E6F288B CRC64;

Query Match 77.5%; Score 31; DB 16; Length 370;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
|||||  
Db 159 VFFAED 164

RESULT 49

Q7W5K7

ID Q7W5K7 PRELIMINARY; PRT; 370 AA.  
AC Q7W5K7;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative integral membrane protein.  
 GN BPP3284.  
 OS Bordetella parapertussis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=12822 / ATCC BAA-587;  
 RX MEDLINE=22827954; PubMed=12910271;  
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL; BX640433; CAE38569.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 370 AA; 40110 MW; C9E4D8073E60823B CRC64;  
  
 Query Match 77.5%; Score 31; DB 16; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 3 VFFAED 8  
 |||||  
 Db 159 VFFAED 164

# RESULT 50

Q7VW49  
 ID Q7VW49 PRELIMINARY; PRT; 370 AA.  
 AC Q7VW49;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative integral membrane protein.  
 GN BP2420.  
 OS Bordetella pertussis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=520;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;  
 RX MEDLINE=22827954; PubMed=12910271;  
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,  
 RA Rabbिनowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,  
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL; BX640418; CAE42691.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 370 AA; 40007 MW; BF3FB5A28484863F CRC64;  
  
 Query Match 77.5%; Score 31; DB 16; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
 |||||  
 Db 159 VFFAED 164

# RESULT 51

O57394  
 ID O57394 PRELIMINARY; PRT; 699 AA.  
 AC O57394;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE EL amyloid precursor protein 699.  
 GN EL APP699.  
 OS Narke japonica (Electric ray).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Squalia; Hypnosqualia; Pristioraja; Batoidea;  
 OC Torpediniformes; Narcinoidei; Narkidae; Narke.  
 OX NCBI\_TaxID=62965;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Electric lobe;  
 RX MEDLINE=98129705; PubMed=9461486;  
 RA Iijima K., Lee D.-S., Okutsu J., Tomita S., Hirashima N., Kirino Y.,  
 RA Suzuki T.;  
 RT "cDNA isolation of Alzheimer's amyloid precursor protein from  
 RT cholinergic nerve terminals of the electric organ of the electric  
 RT ray.";  
 RL Biochem. J. 330:29-33(1998).  
 DR EMBL; AB005544; BAA24230.1; -.  
 DR HSSP; P05067; 1HZ3.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR008155; A4\_APP.  
 DR InterPro; IPR008154; A4\_extra.  
 DR InterPro; IPR001255; Beta-APP.  
 DR Pfam; PF02177; A4\_EXTRA; 1.  
 DR Pfam; PF03494; Beta-APP; 1.  
 DR PRINTS; PR00203; AMYLOIDA4.

DR SMART; SM00006; A4\_EXTRA; 1.  
DR PROSITE; PS00319; A4\_EXTRA; 1.  
DR PROSITE; PS00320; A4\_INTRA; 1.  
SQ SEQUENCE 699 AA; 78879 MW; 952915C309D50E5C CRC64;

Query Match 77.5%; Score 31; DB 13; Length 699;  
Best Local Similarity 75.0%; Pred. No. 4.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||||| :|  
Db 616 KLVFFPKD 623

RESULT 52

Q83E69

ID Q83E69 PRELIMINARY; PRT; 884 AA.  
AC Q83E69;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Pyruvate dehydrogenase, E1 component.  
GN PDHA OR CBU0461.  
OS Coxiella burnetii.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;  
OC Coxiellaceae; Coxiella.  
OX NCBI\_TaxID=777;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nine Mile phase I / RSA 493;  
RX MEDLINE=22608657; PubMed=12704232;  
RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,  
RA Nelson W.C., Ward N.L., Tettelin H., Davidsen T.M., Beanan M.J.,  
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,  
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,  
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;  
RT "Complete genome sequence of the Q-fever pathogen, Coxiella  
RT burnetii.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).  
DR EMBL; AE016961; AAO90011.1; -.  
DR TIGR; CBU0461; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR004660; AceE.  
DR InterPro; IPR005829; Sug\_transporter.  
DR InterPro; IPR005474; Transketolase\_N.  
DR InterPro; IPR009014; Transketo\_C\_like.  
DR Pfam; PF00456; transketolase; 1.  
DR TIGRFAMs; TIGR00759; aceE; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Pyruvate; Complete proteome.  
SQ SEQUENCE 884 AA; 99777 MW; E806FBA03AFA7327 CRC64;

Query Match 77.5%; Score 31; DB 16; Length 884;  
Best Local Similarity 62.5%; Pred. No. 5.3e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||:|: ||  
Db 549 KLIFYRED 556

RESULT 53

Q8AA21

ID Q8AA21 PRELIMINARY; PRT; 906 AA.  
AC Q8AA21;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Pyruvate phosphate dikinase.  
GN BT0644.  
OS Bacteroides thetaiotaomicron.  
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
OC Bacteroidaceae; Bacteroides.  
OX NCBI\_TaxID=818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VPI-5482 / ATCC 29148;  
RX MEDLINE=22550858; PubMed=12663928;  
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
RA Chiang H.C., Hooper L.V., Gordon J.I.;  
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."  
RL Science 299:2074-2076(2003).  
DR EMBL; AE016928; AAO75751.1; -.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0016301; F:kinase activity; IEA.  
DR GO; GO:0016772; F:transferase activity, transferring phosphor. . .; IEA.  
DR GO; GO:0016310; P:phosphorylation; IEA.  
DR InterPro; IPR008279; PEP\_mobile.  
DR InterPro; IPR000121; PEP\_utilizers.  
DR InterPro; IPR002192; PPK\_N\_term.  
DR Pfam; PF00391; PEP-utilizers; 1.  
DR Pfam; PF02896; PEP-utilizers\_C; 1.  
DR Pfam; PF01326; PPK\_N\_term; 1.  
DR ProDom; PD000940; PEP\_utilizers; 1.  
DR PROSITE; PS00742; PEP\_ENZYMES\_2; 1.  
DR PROSITE; PS00370; PEP\_ENZYMES\_PHOS\_SITE; 1.  
KW Kinase; Complete proteome.  
SQ SEQUENCE 906 AA; 99683 MW; FC7A55274E6C2044 CRC64;

Query Match 77.5%; Score 31; DB 16; Length 906;  
Best Local Similarity 62.5%; Pred. No. 5.5e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
::||||:|  
Db 440 QIVFFADD 447

RESULT 54

Q9UUG9

ID Q9UUG9 PRELIMINARY; PRT; 1339 AA.

AC Q9UUG9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative GTPase activating protein, Rap\_GAP domain.  
 GN SPAC630.13C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972h-;  
 RA McDougall R.C., Rajandream M.A., Barrell B.G., Zimmermann W.,  
 RA Wambutt R.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AL109832; CAB52735.1; -.  
 DR PIR; T38991; T38991.  
 DR GeneDB\_SPombe; SPAC630.13c; -.  
 DR GO; GO:0005096; F:GTPase activator activity; IEA.  
 DR GO; GO:0008151; P:cell growth and/or maintenance; IEA.  
 DR InterPro; IPR000331; Rap\_GAP.  
 DR InterPro; IPR003913; Tuberin.  
 DR Pfam; PF02145; Rap\_GAP; 1.  
 DR Pfam; PF03542; Tuberin; 1.  
 DR PROSITE; PS50085; RAPGAP; 1.  
 SQ SEQUENCE 1339 AA; 154812 MW; B4D32AE6F0D100FE CRC64;

Query Match 77.5%; Score 31; DB 3; Length 1339;  
 Best Local Similarity 75.0%; Pred. No. 8e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||||| :|  
 Db 419 KLVFFLKD 426

#### RESULT 55

##### Q8NIT0

ID Q8NIT0 PRELIMINARY; PRT; 1364 AA.  
 AC Q8NIT0;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Probable xanthine dehydrogenase.  
 GN B23E9.040.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [2]



RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AL807372; CAD37030.1; -.  
 DR PIR; T51920; T51920.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR002888; 2Fe-2S\_bind.  
 DR InterPro; IPR006058; 2Fe2S\_fd\_BS.  
 DR InterPro; IPR008274; Aldxan\_dh\_bind.  
 DR InterPro; IPR000674; Aldxan\_dh\_hamm.  
 DR InterPro; IPR005107; CO\_deh\_flav\_C.  
 DR InterPro; IPR002346; dehydrog\_molyb.  
 DR InterPro; IPR001041; Ferredoxin.  
 DR Pfam; PF01315; Ald\_Xan\_dh\_C; 1.  
 DR Pfam; PF02738; Ald\_Xan\_dh\_C2; 1.  
 DR Pfam; PF03450; CO\_deh\_flav\_C; 1.  
 DR Pfam; PF00941; FAD\_binding\_5; 1.  
 DR Pfam; PF00111; fer2; 1.  
 DR Pfam; PF01799; fer2\_2; 1.  
 DR ProDom; PD186071; 2Fe-2S\_bind; 1.  
 DR PROSITE; PS00197; 2FE2S\_FERREDOXIN; 1.  
 KW Iron; Iron-sulfur.  
 SQ SEQUENCE 1364 AA; 149832 MW; 282B3187A22D2127 CRC64;

Query Match 77.5%; Score 31; DB 3; Length 1364;  
 Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
 |||||  
 Db 681 VFFAED 686

# RESULT 56

Q8H8K8

ID Q8H8K8 PRELIMINARY; PRT; 92 AA.  
 AC Q8H8K8;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN OSJNB0078C13.7.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,  
 RA Currie J., Collura K.;  
 RT "Rice Genomic Sequence."  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;  
 RA The Rice Chromosome 10 Sequencing Consortium;  
 RT "In-depth view of structure, activity, and evolution of rice  
 RT chromosome 10.";  
 RL Science 300:1566-1569(2003).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AC091666; AAN05366.1; -.  
 DR EMBL; AE017081; AAP53184.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 92 AA; 9737 MW; 94AE531012888F7D CRC64;

Query Match 75.0%; Score 30; DB 10; Length 92;  
 Best Local Similarity 75.0%; Pred. No. 94;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 || || ||  
 Db 71 KLFFFPED 78

# RESULT 57

O77630

ID O77630 PRELIMINARY; PRT; 122 AA.  
 AC O77630;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Matrix metalloproteinase 3 (Fragment).  
 GN MMP-3.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99357011; PubMed=10429942;  
 RA Flannery C.R., Little C.B., Caterson B., Hughes C.E.;  
 RT "Effects of culture conditions and exposure to catabolic stimulators  
 RT (IL-1 and retinoic acid) on the expression of matrix  
 RT metalloproteinases (MMPs) and disintegrin metalloproteinases (ADAMs)  
 RT by articular cartilage chondrocytes.";  
 RL Matrix Biol. 18:225-237(1999).  
 DR EMBL; AF069642; AAC23526.1; -.  
 DR HSSP; P21692; 1FBL.  
 DR InterPro; IPR000585; Hemopexin.  
 DR Pfam; PF00045; hemopexin; 2.  
 DR SMART; SM00120; HX; 2.  
 DR PROSITE; PS00024; HEMOPEXIN; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 122 122  
 SQ SEQUENCE 122 AA; 14142 MW; E9E33018869B0151 CRC64;

Query Match 75.0%; Score 30; DB 6; Length 122;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| |||||  
Db 74 KTYFFAED 81

RESULT 58

Q8CSR5

ID Q8CSR5 PRELIMINARY; PRT; 125 AA.  
AC Q8CSR5;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Conserved hypothetical protein.  
GN SE0992.  
OS Staphylococcus epidermidis.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 12228;  
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,  
RA Chen Z., Wen Y.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AE016747; AAO04589.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 125 AA; 14741 MW; 0A4AB4D0F3C62B2D CRC64;

Query Match 75.0%; Score 30; DB 16; Length 125;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|| ||||:  
Db 12 KLQFFAEE 19

RESULT 59

Q931Y2

ID Q931Y2 PRELIMINARY; PRT; 204 AA.  
AC Q931Y2;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein SAV0890.  
GN SAV0890.  
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=158878;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21311952; PubMed=11418146;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,

RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
 RT "Whole genome sequencing of meticillin-resistant *Staphylococcus*  
 RT aureus.";  
 RL Lancet 357:1225-1240(2001).  
 DR EMBL; AP003360; BAB57052.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 204 AA; 23664 MW; FE5A54FD38AEE486 CRC64;

Query Match 75.0%; Score 30; DB 16; Length 204;  
 Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 || ||||:  
 Db 10 KLQFFAEE 17

# RESULT 60

Q21915

ID Q21915 PRELIMINARY; PRT; 222 AA.  
 AC Q21915;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE R10H10.1 protein.  
 GN R10H10.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bardill S.C.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode *C.elegans*: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z70686; CAA94609.1; -.  
 DR PIR; T24151; T24151.  
 DR WormPep; R10H10.1; CE06292.  
 DR InterPro; IPR001075; NifU\_C.  
 DR Pfam; PF01106; NifU-like; 1.  
 DR ProDom; PD002830; NifU\_C; 1.  
 SQ SEQUENCE 222 AA; 24757 MW; ED232B974B03E8D0 CRC64;

Query Match 75.0%; Score 30; DB 5; Length 222;  
 Best Local Similarity 75.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| ||| ||.  
Db 69 KRVFFGED 76

RESULT 61

Q8ZAM4

ID Q8ZAM4 PRELIMINARY; PRT; 258 AA.  
AC Q8ZAM4;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Sec-independent protein translocase protein TatC (Hypothetical  
DE protein).  
GN TATC OR YPO3776 OR Y0454.  
OS *Yersinia pestis*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; *Yersinia*.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RX MEDLINE=21470413; PubMed=11586360;  
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
RT "Genome sequence of *Yersinia pestis*, the causative agent of plague.";  
RL Nature 413:523-527(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KIM5 / Biovar Mediaevalis;  
RX MEDLINE=22137863; PubMed=12142430;  
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of *Yersinia pestis* KIM.";  
RL J. Bacteriol. 184:4601-4611(2002).  
DR EMBL; AJ414158; CAC93243.1; -.  
DR EMBL; AE013646; AAM84043.1; -.  
DR PIR; AG0459; AG0459.  
DR InterPro; IPR008277; Tat\_C.  
DR InterPro; IPR002033; Translocase.  
DR Pfam; PF00902; TatC; 1.  
DR TIGRFAMs; TIGR00945; tatC; 1.  
DR PROSITE; PS01218; TATC; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 258 AA; 28725 MW; 7B052D8306E6FDA6 CRC64;

Query Match 75.0%; Score 30; DB 16; Length 258;  
Best Local Similarity 85.7%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LVFFAED 8  
| | | | |  
Db 34 LVFFAND 40

RESULT 62

Q898W5

ID Q898W5 PRELIMINARY; PRT; 354 AA.  
AC Q898W5;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Conserved protein.  
GN CTC00323.  
OS Clostridium tetani.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Massachusetts / E88;  
RX MEDLINE=22457253; PubMed=12552129;  
RA Brueggemann H., Baeumer S., Fricke W.F., Wiezer A., Liesegang H.,  
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,  
RA Gottschalk G.;  
RT "The genome sequence of Clostridium tetani, the causative agent of  
RT tetanus disease."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).  
DR EMBL; AE015937; AAO34964.1; -.  
KW Complete proteome.  
SQ SEQUENCE 354 AA; 42087 MW; ED52A7D1C86C95BC CRC64;

Query Match 75.0%; Score 30; DB 16; Length 354;  
Best Local Similarity 71.4%; Pred. No. 3.5e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LVFFAED 8  
: | | | | :  
Db 289 MVFFADD 295

RESULT 63

Q93CU7

ID Q93CU7 PRELIMINARY; PRT; 364 AA.  
AC Q93CU7;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Putative glycosyl transferase.  
GN WBDG.  
OS Shigella boydii.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Shigella.  
OX NCBI\_TaxID=621;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=21481970; PubMed=11598067;  
 RA Wang L., Qu W., Reeves P.R.;  
 RT "Sequence Analysis of Four Shigella boydii O-Antigen Loci: Implication  
 RT for Escherichia coli and Shigella Relationships.";  
 RL Infect. Immun. 69:6923-6930(2001).  
 DR EMBL; AF402312; AAL27317.1; -.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR InterPro; IPR001296; Glyco\_trans\_1.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 KW Transferase.  
 SQ SEQUENCE 364 AA; 42395 MW; 3A4B7684A85DEEF7 CRC64;

Query Match 75.0%; Score 30; DB 2; Length 364;  
 Best Local Similarity 62.5%; Pred. No. 3.6e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||:| :|  
 Db 119 KLVYFVQD 126

#### RESULT 64

##### Q8NEN2

ID Q8NEN2 PRELIMINARY; PRT; 370 AA.  
 AC Q8NEN2;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Similar to G protein-coupled receptor 85.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA Strausberg R.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC030577; AAH30577.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPSN.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 370 AA; 41964 MW; 20DD032E716BC797 CRC64;

Query Match 75.0%; Score 30; DB 4; Length 370;  
 Best Local Similarity 62.5%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||:|| |

Db 204 KLIFVHD 211

RESULT 65

Q8ZNH0

ID Q8ZNH0 PRELIMINARY; PRT; 441 AA.  
AC Q8ZNH0;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative permease.  
GN STM2274.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
RT LT2.";  
RL Nature 413:852-856(2001).  
DR EMBL; AE008802; AAL21175.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR001064; Crystallin.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.  
DR Pfam; PF00083; sugar\_tr; 1.  
DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
DR PROSITE; PS50850; MFS; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 441 AA; 48972 MW; D03488ADD17F22FB CRC64;

Query Match 75.0%; Score 30; DB 16; Length 441;  
Best Local Similarity 71.4%; Pred. No. 4.4e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
|:|:|:|  
Db 99 LIFFAKD 105

RESULT 66

O62511

ID O62511 PRELIMINARY; PRT; 502 AA.  
AC O62511;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)



DE ZK550.2 protein.  
 GN ZK550.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Basham V.M.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z82287; CAB05312.1; -.  
 DR PIR; T27908; T27908.  
 DR WormPep; ZK550.2; CE16733.  
 DR InterPro; IPR007114; MFS.  
 DR PROSITE; PS50850; MFS; 1.  
 SQ SEQUENCE 502 AA; 55666 MW; B1B98A5B34254056 CRC64;

Query Match 75.0%; Score 30; DB 5; Length 502;  
 Best Local Similarity 85.7%; Pred. No. 5e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LVFFAED 8  
 |||| ||  
 Db 227 LVFFQED 233

# RESULT 67

Q8A226

ID Q8A226 PRELIMINARY; PRT; 513 AA.  
 AC Q8A226;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Conserved hypothetical protein.  
 GN BT3482.  
 OS Bacteroides thetaiotaomicron.  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Bacteroidaceae; Bacteroides.  
 OX NCBI\_TaxID=818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VPI-5482 / ATCC 29148;  
 RX MEDLINE=22550858; PubMed=12663928;  
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
 RA Chiang H.C., Hooper L.V., Gordon J.I.;  
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."  
 RL Science 299:2074-2076(2003).  
 DR EMBL; AE016940; AAO78588.1; -.  
 DR GO; GO:0030089; C:phycobilisome; IEA.  
 DR GO; GO:0015979; P:photosynthesis; IEA.

DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR002909; IPT\_TIG.  
 DR InterPro; IPR001258; NHL.  
 DR InterPro; IPR001297; PBS\_linker\_poly.  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR Pfam; PF01436; NHL; 1.  
 DR Pfam; PF01833; TIG; 1.  
 DR ProDom; PD000852; PBS\_linker\_poly; 1.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 513 AA; 57650 MW; 79055DC7027D35FE CRC64;

Query Match 75.0%; Score 30; DB 16; Length 513;  
 Best Local Similarity 62.5%; Pred. No. 5.1e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 :|::|  
 Db 187 ELLYFAED 194

# RESULT 68

Q8Q077

ID Q8Q077 PRELIMINARY; PRT; 607 AA.  
 AC Q8Q077;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Putative sensory transduction protein kinase (EC 2.7.3.-).  
 GN MM0260.  
 OS Methanosarcina mazei (Methanosarcina frisia).  
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;  
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.  
 OX NCBI\_TaxID=2209;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;  
 RX MEDLINE=22120827; PubMed=12125824;  
 RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,  
 RA Martinez-Arias R., Henne A., Wiezer A., Baeumer S., Jacobi C.,  
 RA Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,  
 RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
 RA Fritz H.-J., Gottschalk G.;  
 RT "The genome of Methanosarcina mazei: evidence for lateral gene  
 RT transfer between Bacteria and Archaea."  
 RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
 DR EMBL; AE013250; AAM29956.1; -.  
 DR GO; GO:0016301; F:kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.  
 DR InterPro; IPR003018; GAF.  
 DR InterPro; IPR001610; PAC.  
 DR InterPro; IPR000700; PAS-assoc\_C.  
 DR InterPro; IPR000014; PAS\_domain.  
 DR Pfam; PF01590; GAF; 1.  
 DR Pfam; PF00785; PAC; 3.

DR TIGRFAMs; TIGR00229; sensory\_box; 3.  
DR PROSITE; PS50113; PAC; 3.  
DR PROSITE; PS50112; PAS; 1.  
KW Kinase; Transferase; Complete proteome; Hypothetical protein.  
SQ SEQUENCE 607 AA; 69922 MW; F2CB8A10D23D53AF CRC64;

Query Match 75.0%; Score 30; DB 17; Length 607;  
Best Local Similarity 75.0%; Pred. No. 6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|| ||||:  
Db 533 KLSFFAEE 540

RESULT 69

O60157

ID O60157 PRELIMINARY; PRT; 681 AA.  
AC O60157;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Conserved hypothetical protein.  
GN SPBC19C7.08C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972h-;  
RA Lyne M., Rajandream M.A., Barrell B.G., Lucas M., Gaillardin C.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AL023859; CAA19576.1; -.  
DR PIR; T39814; T39814.  
DR GeneDB\_SPombe; SPBC19C7.08c; -.  
DR InterPro; IPR007213; LCM.  
DR Pfam; PF04072; LCM; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 681 AA; 77552 MW; 25070B74822D39F8 CRC64;

Query Match 75.0%; Score 30; DB 3; Length 681;  
Best Local Similarity 62.5%; Pred. No. 6.8e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|:||| :|  
Db 589 KVVFFGDD 596

RESULT 70

Q03455

ID Q03455 PRELIMINARY; PRT; 740 AA.  
AC Q03455;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN MSC2 OR YD8142.02 OR YDR205W.  
 OS *Saccharomyces cerevisiae* (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB972;  
 RA Oliver K., Shore L., Harris D.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB972;  
 RA Barrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; Z68194; CAA92344.1; -.  
 DR PIR; S61568; S61568.  
 DR SGD; S0002613; MSC2.  
 DR GO; GO:0005783; C:endoplasmic reticulum; IDA.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0015491; F:cation:antiporter activity; IMP.  
 DR GO; GO:0006882; P:zinc ion homeostasis; IMP.  
 DR InterPro; IPR002524; Cation\_efflux.  
 DR Pfam; PF01545; Cation\_efflux; 1.  
 DR TIGRFAMs; TIGR01297; CDF; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 740 AA; 82471 MW; E517D782B8CA3986 CRC64;

Query Match 75.0%; Score 30; DB 3; Length 740;  
 Best Local Similarity 62.5%; Pred. No. 7.3e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 ||:||: |  
 Db 65 KLIFSTD 72

# RESULT 71

O77066

ID O77066 PRELIMINARY; PRT; 748 AA.  
 AC O77066;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Diapause associated protein 2.  
 GN DAP2.  
 OS *Choristoneura fumiferana* (Spruce budworm).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Tortricidae; Tortricinae; Choristoneura.  
 OX NCBI\_TaxID=7141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Palli S.R., Ladd T.R., Ricci A.R., Primavera M., Mungrue I.N.,  
 RA Pang A.S.D., Retnakaran A.;

RT "Synthesis of the same two proteins prior to larval diapause and  
RT pupation in the spruce budworm, *Choristoneura fumiferana*.";  
RL J. Insect Physiol. 44:509-524(1998).  
DR EMBL; AF007768; AAC35429.1; -.  
DR HSSP; P04253; 1LLA.  
DR GO; GO:0005344; F:oxygen transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR008922; Di-copper\_centre.  
DR InterPro; IPR000896; Hemocyanin.  
DR InterPro; IPR005203; hemocyanin\_C.  
DR InterPro; IPR005204; hemocyanin\_N.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00372; hemocyanin; 1.  
DR Pfam; PF03723; hemocyanin\_C; 1.  
DR Pfam; PF03722; hemocyanin\_N; 1.  
DR PRINTS; PR00187; HAEMOCYANIN.  
DR PROSITE; PS00209; HEMOCYANIN\_1; 1.  
DR PROSITE; PS00210; HEMOCYANIN\_2; 1.  
SQ SEQUENCE 748 AA; 88192 MW; 7354407BCF285CF5 CRC64;

Query Match 75.0%; Score 30; DB 5; Length 748;  
Best Local Similarity 62.5%; Pred. No. 7.4e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
||::||:|  
Db 108 KLLYFAKD 115

# RESULT 72

## Q9LFI1

ID Q9LFI1 PRELIMINARY; PRT; 768 AA.  
AC Q9LFI1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN F4P12\_60.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bloecker H., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,  
RA Salanoubat M.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AL132966; CAB67643.1; -.  
DR PIR; T45876; T45876.  
DR InterPro; IPR002885; PPR.  
DR Pfam; PF01535; PPR; 13.  
DR TIGRFAMs; TIGR00756; PPR; 9.

KW Hypothetical protein.

SQ SEQUENCE 768 AA; 85674 MW; 5A9060E614B44EE9 CRC64;

Query Match 75.0%; Score 30; DB 10; Length 768;  
Best Local Similarity 83.3%; Pred. No. 7.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VFFAED 8  
:|||||  
Db 724 IFFAED 729

RESULT 73

Q9P852

ID Q9P852 PRELIMINARY; PRT; 810 AA.  
AC Q9P852;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Vps18 protein.  
GN VPS18.  
OS Candida albicans (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
OX NCBI\_TaxID=5476;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Palmer G.E., Cashmore A.;  
RT "Characterisation of a Candida albicans VPS18 homologue."  
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ289080; CAB91628.1; -.  
DR InterPro; IPR000547; Clathrin\_repeat.  
DR InterPro; IPR007810; Pep3\_Vps18.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00637; Clathrin; 1.  
DR Pfam; PF05131; Pep3\_Vps18; 1.  
DR SMART; SM00184; RING; 1.  
SQ SEQUENCE 810 AA; 93887 MW; AF637CFCD4527CA3 CRC64;

Query Match 75.0%; Score 30; DB 3; Length 810;  
Best Local Similarity 71.4%; Pred. No. 8e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAE 7  
||:|:|  
Db 467 KLIFYAE 473

RESULT 74

Q97S58

ID Q97S58 PRELIMINARY; PRT; 99 AA.  
AC Q97S58;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Ribosomal protein L7A family.  
GN SP0555.

OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC BAA-334 / TIGR4;  
 RX MEDLINE=21357209; PubMed=11463916;  
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,  
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,  
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
 RT "Complete genome sequence of a virulent isolate of Streptococcus  
 RT pneumoniae.";  
 RL Science 293:498-506(2001).  
 DR EMBL; AE007365; AAK74711.1; -.  
 DR PIR; F95064; F95064.  
 DR TIGR; SP0555; -.  
 DR GO; GO:0005840; C:ribosome; IEA.  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 DR GO; GO:0006412; P:protein biosynthesis; IEA.  
 DR InterPro; IPR004038; Ribosomal\_L7A.  
 DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 99 AA; 10886 MW; 23E4A396EB046DFB CRC64;

Query Match 72.5%; Score 29; DB 16; Length 99;  
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |||| | |  
 Db 36 KLVFLAHD 43

# RESULT 75

Q8DQV3

ID Q8DQV3 PRELIMINARY; PRT; 99 AA.  
 AC Q8DQV3;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Conserved hypothetical protein.  
 GN SPR0480.  
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=171101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21429245; PubMed=11544234;  
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszczyk L.C., Burgett S.,  
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,

RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,  
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,  
 RA McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I.,  
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,  
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,  
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rosteck P.R. Jr., Skatrud P.L.,  
 RA Glass J.I.;  
 RT "Genome of the bacterium *Streptococcus pneumoniae* strain R6.";  
 RL J. Bacteriol. 183:5709-5717(2001).  
 DR EMBL; AE008427; AAK99284.1; -.  
 DR PIR; H97931; H97931.  
 DR GO; GO:0005840; C:ribosome; IEA.  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 DR GO; GO:0006412; P:protein biosynthesis; IEA.  
 DR InterPro; IPR004038; Ribosomal\_L7A.  
 DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 99 AA; 10902 MW; 3FF8A396EB112DFB CRC64;

Query Match 72.5%; Score 29; DB 16; Length 99;  
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |||| |  
 Db 36 KLVFLAHD 43

# RESULT 76

Q8Y8G0

ID Q8Y8G0 PRELIMINARY; PRT; 100 AA.  
 AC Q8Y8G0;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein lmo0944.  
 GN LMO0944.  
 OS *Listeria monocytogenes*.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD-e / Serovar 1/2a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of *Listeria* species.";  
 RL Science 294:849-852(2001).



DR EMBL; AL591977; CAC99022.1; -.  
DR ListiList; LMO00944; -.  
DR InterPro; IPR008326; UCP034852.  
DR PIRSF; PIRSF034852; UCP034852; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 100 AA; 11693 MW; 6BCCD399F9417AC7 CRC64;

Query Match 72.5%; Score 29; DB 16; Length 100;  
Best Local Similarity 71.4%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVFFAED 8  
|:| |  
Db 63 LIFFIED 69

#### RESULT 77

Q88VK6

ID Q88VK6 PRELIMINARY; PRT; 101 AA.  
AC Q88VK6;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Ribosomal protein HS6-type (S12/L30/L7a).  
GN LP\_2041.  
OS Lactobacillus plantarum.  
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
OC Lactobacillus.  
OX NCBI\_TaxID=1590;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCIMB 8826 / WCFS1;  
RX MEDLINE=22480296; PubMed=12566566;  
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,  
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,  
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,  
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,  
RA De Vos W.M., Siezen R.J.;  
RT "Complete genome sequence of Lactobacillus plantarum WCFS1."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).  
DR EMBL; AL935257; CAD64415.1; -.  
DR GO; GO:0005840; C:ribosome; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
DR GO; GO:0006412; P:protein biosynthesis; IEA.  
DR InterPro; IPR004038; Ribosomal\_L7A.  
DR Pfam; PF01248; Ribosomal\_L7Ae; 1.  
KW Complete proteome.  
SQ SEQUENCE 101 AA; 10786 MW; D36759BAA6509956 CRC64;

Query Match 72.5%; Score 29; DB 16; Length 101;  
Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
| | | | |  
Db 37 KLVFIASD 44

RESULT 78

Q9GMZ2

ID Q9GMZ2 PRELIMINARY; PRT; 119 AA.  
 AC Q9GMZ2;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Voltage-dependent potassium channel (Fragment).  
 GN KV2.2.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ohya S., Imaizumi Y.;  
 RT "porcine voltage-dependent potassium channel (Kv2.2).";  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AB042526; BAB12398.1; -.  
 DR HSSP; Q54397; 1BL8.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0008076; C:voltage-gated potassium channel complex; IEA.  
 DR GO; GO:0005249; F:voltage-gated potassium channel activity; IEA.  
 DR GO; GO:0006812; P:cation transport; IEA.  
 DR GO; GO:0006813; P:potassium ion transport; IEA.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR001622; K+channel\_pore.  
 DR InterPro; IPR003968; Kv\_channel.  
 DR InterPro; IPR003091; K\_channel.  
 DR InterPro; IPR005820; M+channel\_nlg.  
 DR InterPro; IPR003973; Shab\_channel.  
 DR Pfam; PF00520; ion\_trans; 1.  
 DR PRINTS; PR00169; KCHANNEL.  
 DR PRINTS; PR01491; KVCHANNEL.  
 DR PRINTS; PR01495; SHABCHANNEL.  
 KW Ionic channel; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 119 119  
 SQ SEQUENCE 119 AA; 13382 MW; EA856E73FAA6CE64 CRC64;

Query Match 72.5%; Score 29; DB 6; Length 119;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVFFAE 7  
 |||||  
 Db 25 LVFFAE 30

RESULT 79

Q7UPR1

ID Q7UPR1 PRELIMINARY; PRT; 137 AA.  
 AC Q7UPR1;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.  
 GN RB6777.  
 OS Rhodopirellula baltica.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
 RA Schlesner H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 RT strain 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
 DR EMBL; BX294144; CAD75000.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 137 AA; 15530 MW; 5DCC133B06CC5FC0 CRC64;

Query Match 72.5%; Score 29; DB 16; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFA 6  
 |||||  
 Db 85 KLVFFA 90

#### RESULT 80

##### Q9STZ9

ID Q9STZ9 PRELIMINARY; PRT; 152 AA.  
 AC Q9STZ9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN T20K18.220 OR AT4G12870.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,  
 RA Bancroft I., Mewes H.W., Mayer K.F.X., Schueller C.;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,  
 RA Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.

RN [4]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AL049640; CAB41004.1; -.  
 DR EMBL; AL161535; CAB78329.1; -.  
 DR PIR; T06645; T06645.  
 DR InterPro; IPR004911; GILT.  
 DR Pfam; PF03227; GILT; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 152 AA; 17095 MW; ED47CEAE7607B131 CRC64;

Query Match 72.5%; Score 29; DB 10; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFA 6  
 |||||  
 Db 9 KLVFFA 14

# RESULT 81

Q32197

ID Q32197 PRELIMINARY; PRT; 160 AA.  
 AC Q32197;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Eichhornia meyeri.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Eichhornia.  
 OX NCBI\_TaxID=44950;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae."  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41602; AAD09695.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 160 160  
 SQ SEQUENCE 160 AA; 18517 MW; 8302150D22A72F5B CRC64;

Query Match 72.5%; Score 29; DB 8; Length 160;  
 Best Local Similarity 50.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 82

Q32892

ID Q32892 PRELIMINARY; PRT; 162 AA.  
AC Q32892;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NADH dehydrogenase subunit F (Fragment).  
GN NDHF.  
OS Pontederia rotundifolia.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
OC Pontederiaceae; Pontederia.  
OX NCBI\_TaxID=44963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22062003; PubMed=12066301;  
RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
RT "Phylogenetic congruence and discordance among one morphological and  
RT three molecular data sets from Pontederiaceae."  
RL Syst. Biol. 47:545-567(1998).  
DR EMBL; U41620; AAD09711.1; -.  
DR PIR; T13656; T13656.  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
DR InterPro; IPR002128; Oxidored\_q1\_C.  
DR Pfam; PF01010; oxidored\_q1\_C; 1.  
KW Chloroplast.  
FT NON\_TER 1 1  
FT NON\_TER 162 162  
SQ SEQUENCE 162 AA; 18720 MW; 2C0081C81D8E2638 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 83

Q32202

ID Q32202 PRELIMINARY; PRT; 162 AA.  
AC Q32202;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NADH dehydrogenase subunit F (Fragment).

GN NDHF.  
 OS Eichhornia paradoxa.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Eichhornia.  
 OX NCBI\_TaxID=44952;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae."  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41607; AAD09700.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18768 MW; 3EBF8733F3200AFE CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |::||::|  
 Db 13 KVIFFSKD 20

#### RESULT 84

Q32176

ID Q32176 PRELIMINARY; PRT; 162 AA.  
 AC Q32176;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Eichhornia crassipes (Water hyacinth).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Eichhornia.  
 OX NCBI\_TaxID=44947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae."  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41599; AAD09842.1; -.

DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18598 MW; C959E914F97AC427 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |::||::|  
 Db 13 KVIFFSKD 20

# RESULT 85

Q32430

ID Q32430 PRELIMINARY; PRT; 162 AA.  
 AC Q32430;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Heteranthera seubertiana.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Heteranthera.  
 OX NCBI\_TaxID=44961;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae."  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41611; AAD09702.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18503 MW; 229586412B990FD2 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8

Db                   |::||::|  
13 KVIFFSKD 20

RESULT 86

Q32201

ID   Q32201           PRELIMINARY;       PRT;   162 AA.  
AC   Q32201;  
DT   01-NOV-1996 (TrEMBLrel. 01, Created)  
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE   NADH dehydrogenase subunit F (Fragment).  
GN   NDHF.  
OS   Eichhornia paniculata.  
OG   Chloroplast.  
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC   Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
OC   Pontederiaceae; Eichhornia.  
OX   NCBI\_TaxID=44951;  
RN   [1]  
RP   SEQUENCE FROM N.A.  
RX   MEDLINE=22062003; PubMed=12066301;  
RA   Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
RT   "Phylogenetic congruence and discordance among one morphological and  
RT   three molecular data sets from Pontederiaceae."  
RL   Syst. Biol. 47:545-567(1998).  
DR   EMBL; U41603; AAD09696.1; -.  
DR   GO; GO:0009507; C:chloroplast; IEA.  
DR   GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
DR   GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
DR   InterPro; IPR002128; Oxidored\_q1\_C.  
DR   Pfam; PF01010; oxidored\_q1\_C; 1.  
KW   Chloroplast.  
FT   NON\_TER           1       1  
FT   NON\_TER       162       162  
SQ   SEQUENCE   162 AA; 18647 MW; 372BC625DEEAD127 CRC64;

Query Match                   72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity   50.0%; Pred. No. 2.7e+02;  
Matches       4; Conservative   4; Mismatches       0; Indels       0; Gaps       0;

Qy           1 KLVFFAED 8  
             |::||::|  
Db           13 KVIFFSKD 20

RESULT 87

Q32441

ID   Q32441           PRELIMINARY;       PRT;   162 AA.  
AC   Q32441;  
DT   01-NOV-1996 (TrEMBLrel. 01, Created)  
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE   NADH dehydrogenase subunit F (Fragment).  
GN   NDHF.  
OS   Heteranthera zosterifolia.  
OG   Chloroplast.



OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Heteranthera.  
 OX NCBI\_TaxID=44962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae."  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41612; AAD09703.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18616 MW; 9FDECF40FA8092F1 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |::||::|  
 Db 13 KVIFFSKD 20

# RESULT 88

Q32802

ID Q32802 PRELIMINARY; PRT; 162 AA.  
 AC Q32802;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Pontederia cordata (Pickerel weed).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Pontederia.  
 OX NCBI\_TaxID=44964;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae."  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41619; AAD09710.1; -.  
 DR PIR; T13563; T13563.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.

DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
DR InterPro; IPR002128; Oxidored\_q1\_C.  
DR Pfam; PF01010; oxidored\_q1\_C; 1.  
KW Chloroplast.  
FT NON\_TER 1 1  
FT NON\_TER 162 162  
SQ SEQUENCE 162 AA; 18686 MW; 2C0081C81D868438 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

# RESULT 89

Q32424

ID Q32424 PRELIMINARY; PRT; 162 AA.  
AC Q32424;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NADH dehydrogenase subunit F (Fragment).  
GN NDHF.  
OS Heteranthera rotundifolia.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
OC Pontederiaceae; Heteranthera.  
OX NCBI\_TaxID=44960;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22062003; PubMed=12066301;  
RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
RT "Phylogenetic congruence and discordance among one morphological and  
RT three molecular data sets from Pontederiaceae."  
RL Syst. Biol. 47:545-567(1998).  
DR EMBL; U41610; AAD09848.1; -.  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
DR InterPro; IPR002128; Oxidored\_q1\_C.  
DR Pfam; PF01010; oxidored\_q1\_C; 1.  
KW Chloroplast.  
FT NON\_TER 1 1  
FT NON\_TER 162 162  
SQ SEQUENCE 162 AA; 18669 MW; 7232BF0201A09858 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 90

Q32556

ID Q32556 PRELIMINARY; PRT; 162 AA.  
 AC Q32556;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Monochoria cyanea.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Monochoria.  
 OX NCBI\_TaxID=44969;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae.";  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41613; AAD09704.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18759 MW; 0C7AA56F8ADA5F29 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |::||::|  
 Db 13 KVIFFSKD 20

RESULT 91

Q32588

ID Q32588 PRELIMINARY; PRT; 162 AA.  
 AC Q32588;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Monochoria korsakowii.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;

OC Pontederiaceae; Monochoria.  
 OX NCBI\_TaxID=44971;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae.";  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41615; AAD09706.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18734 MW; 4071D0CF24A8F3D6 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |::||::|  
 Db 13 KVIFFSKD 20

# RESULT 92

Q32390

ID Q32390 PRELIMINARY; PRT; 162 AA.  
 AC Q32390;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Heteranthera dubia (water star-grass).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Heteranthera.  
 OX NCBI\_TaxID=44955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae.";  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41605; AAD09698.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.

KW Chloroplast.  
FT NON\_TER 1 1  
FT NON\_TER 162 162  
SQ SEQUENCE 162 AA; 18414 MW; 5E68663379B1F63F CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 93

Q32164

ID Q32164 PRELIMINARY; PRT; 162 AA.  
AC Q32164;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NADH dehydrogenase subunit F (Fragment).  
GN NDHF.  
OS Eichhornia azurea.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
OC Pontederiaceae; Eichhornia.  
OX NCBI\_TaxID=44946;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22062003; PubMed=12066301;  
RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
RT "Phylogenetic congruence and discordance among one morphological and  
RT three molecular data sets from Pontederiaceae.";  
RL Syst. Biol. 47:545-567(1998).  
DR EMBL; U41598; AAD09692.1; -.  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
DR InterPro; IPR002128; Oxidored\_q1\_C.  
DR Pfam; PF01010; oxidored\_q1\_C; 1.  
KW Chloroplast.  
FT NON\_TER 1 1  
FT NON\_TER 162 162  
SQ SEQUENCE 162 AA; 18765 MW; 3CE08836AE0370F1 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 94

Q32579

ID Q32579 PRELIMINARY; PRT; 162 AA.  
AC Q32579;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NADH dehydrogenase subunit F (Fragment).  
GN NDHF.  
OS Monochoria hastata.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
OC Pontederiaceae; Monochoria.  
OX NCBI\_TaxID=44970;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22062003; PubMed=12066301;  
RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
RT "Phylogenetic congruence and discordance among one morphological and  
RT three molecular data sets from Pontederiaceae.";  
RL Syst. Biol. 47:545-567(1998).  
DR EMBL; U41614; AAD09705.1; -.  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
DR InterPro; IPR002128; Oxidored\_q1\_C.  
DR Pfam; PF01010; oxidored\_q1\_C; 1.  
KW Chloroplast.  
FT NON\_TER 1 1  
FT NON\_TER 162 162  
SQ SEQUENCE 162 AA; 18783 MW; 4F90D4CF31D2F51B CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 95

Q32190

ID Q32190 PRELIMINARY; PRT; 162 AA.  
AC Q32190;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NADH dehydrogenase subunit F (Fragment).  
GN NDHF.  
OS Eichhornia heterosperma.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
OC Pontederiaceae; Eichhornia.  
OX NCBI\_TaxID=44949;  
RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae.";  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41601; AAD09694.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18765 MW; 3CE08836AE0370F1 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
 |::||::|  
 Db 13 KVIFFSKD 20

# RESULT 96

Q32915

ID Q32915 PRELIMINARY; PRT; 162 AA.  
 AC Q32915;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Pontederia sagittata.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Pontederia.  
 OX NCBI\_TaxID=16373;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae.";  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41621; AAD09712.1; -.  
 DR PIR; T13659; T13659.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1

FT NON\_TER 162 162  
SQ SEQUENCE 162 AA; 18720 MW; C1BBE1C81D868C10 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 97

Q32400

ID Q32400 PRELIMINARY; PRT; 162 AA.  
AC Q32400;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NADH dehydrogenase subunit F (Fragment).  
GN NDHF.  
OS Hydrothrix gardneri.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
OC Pontederiaceae; Hydrothrix.  
OX NCBI\_TaxID=44957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22062003; PubMed=12066301;  
RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
RT "Phylogenetic congruence and discordance among one morphological and  
RT three molecular data sets from Pontederiaceae."  
RL Syst. Biol. 47:545-567(1998).  
DR EMBL; U41606; AAD09699.1; -.  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
DR InterPro; IPR002128; Oxidored\_q1\_C.  
DR Pfam; PF01010; oxidored\_q1\_C; 1.  
KW Chloroplast.  
FT NON\_TER 1 1  
FT NON\_TER 162 162  
SQ SEQUENCE 162 AA; 18692 MW; 83FD22EC18FB68C7 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

RESULT 98

Q32648

ID Q32648 PRELIMINARY; PRT; 162 AA.



AC Q32648;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Monochoria vaginalis.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Monochoria.  
 OX NCBI\_TaxID=44972;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae."  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41616; AAD09707.1; -.  
 DR PIR; T13487; T13487.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18778 MW; 0ADB64D58D04850D CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |::||::|  
 Db 13 KVIFFSKD 20

# RESULT 99

Q32419

ID Q32419 PRELIMINARY; PRT; 162 AA.  
 AC Q32419;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Heteranthera oblongifolia.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Heteranthera.  
 OX NCBI\_TaxID=44959;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae.";  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41609; AAD09701.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18644 MW; 71109BBEA1303972 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFFAED 8  
 |::||::|  
 Db 13 KVIFFSKD 20

# RESULT 100

Q32179

ID Q32179 PRELIMINARY; PRT; 162 AA.  
 AC Q32179;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit F (Fragment).  
 GN NDHF.  
 OS Eichhornia diversifolia.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Commelinales;  
 OC Pontederiaceae; Eichhornia.  
 OX NCBI\_TaxID=44948;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22062003; PubMed=12066301;  
 RA Graham S.W., Kohn J.R., Morton B.R., Eckenwalder J.E., Barrett S.C.;  
 RT "Phylogenetic congruence and discordance among one morphological and  
 RT three molecular data sets from Pontederiaceae.";  
 RL Syst. Biol. 47:545-567(1998).  
 DR EMBL; U41600; AAD09693.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR Pfam; PF01010; oxidored\_q1\_C; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 162 162  
 SQ SEQUENCE 162 AA; 18765 MW; 3CE08836AE0370F1 CRC64;

Query Match 72.5%; Score 29; DB 8; Length 162;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFFAED 8  
|::||::|  
Db 13 KVIFFSKD 20

Search completed: February 28, 2004, 08:55:49  
Job time : 147.5 secs